GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 17, 2003, 07:00:08 ; Search time 14 Seconds
(without alignments)
1318.415 Million cell updates/sec

Title: US-09-915-873-3
Perfect score: 178
Sequence: 1 accatgcagatcctgaacac.....gcagcaaccctcacgcgcgcgtga 96

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 566448

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame_n2p.model DEV xth
-Q=/cgn2_1/PSPT/sp_1_US09_915_873_3.seq -DB=PIR_73 -QFMT=tastan -SUFFIX=n2p.rpr -MINMATCH=0.1 -LOGPCL 0 -LOGPEXT 0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST 45
-DOCALIGN=200 -THR_SCORE pct -THR_MAX 100 -THR_MIN 0 -MATCH_LEN_MIN
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN 2000000000
-USER=US09915873_2CGN_1_1_10_Grmat_14012003_171603_13839 -NCPU 6 -DEBUG
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES 0 -WAIT -LOADCOG -DEV_TIMEOUT 120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP 6 -FGAPEXT 7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP 6 -DELEXT=7

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	146	82.0	434	2	S12775	alpha-amylase (EC
2	117	65.7	428	2	S10013	alpha-amylase (EC
3	104	58.4	383	1	ALRZOC	alpha-amylase (EC
4	71.5	40.2	427	1	ALBHB	alpha-amylase (EC
5	69.5	39.0	429	1	JE0406	alpha-amylase (EC
6	67	37.6	386	2	T36169	hypothetical prote
7	66	37.1	354	2	T49806	hypothetical prote
8	66	37.1	456	2	F88494	protein F57B9.4 [i
c 9	63	37.3	185	2	D71654	FKBP type peptidyl
c 10	63	37.3	185	2	C64660	peptidyl-prolyl ci
11	62.5	35.1	805	2	T49385	hypothetical prote
12	61.5	34.6	698	2	T32594	hypothetical prote
13	60	33.7	892	2	T09193	[alpha 2 - human
14	59.5	33.4	162	2	E87685	hypothetical prote
c 15	59	33.1	209	1	A26A2	[alpha-2A-adr
c 16	59	33.1	209	1	A2AR	[alpha-2A-adr
17	59	33.1	1374	2	T51343	[alpha-2A-adr
c 18	58.5	32.9	209	2	S13413	DNA-binding protein
c 19	58	32.6	380	2	I47500	hypothetical prote
20	57.5	32.3	905	1	I75766	[alpha-2A-adr
c 21	57.5	32.3	935	1	I75767	[alpha-2A-adr
c 22	57.5	32.3	935	1	S16858	[alpha-2A-adr
23	57	32.0	157	2	B72047	hypothetical prote
24	57	32.0	268	2	S41617	protein kinase (EC
25	57	32.0	512	1	B46273	[alpha-2A-adr
26	57	32.0	527	2	B64781	hypothetical prote
27	57	32.0	1385	2	S16291	pol protein - [vir
28	57	32.0	4967	2	I64456	Arp protein - [vir
29	57	32.0	5062	2	I64743	Arp protein - [vir
30	56.5	31.7	84	2	S61143	NA protein - [vir
31	56.5	31.7	596	2	I49766	hypothetical prote
32	56.5	31.7	804	2	S09146	hypothetical prote
33	56	31.5	170	2	B56766	[alpha-2A-adr
34	56	31.5	216	2	I44766	hypothetical prote
35	56	31.5	576	2	A46766	hypothetical prote
36	56	31.5	1240	2	I49766	hypothetical prote
37	55.5	31.2	282	2	I41766	hypothetical prote
c 38	55.5	31.2	344	2	B47766	hypothetical prote
c 39	55.5	31.2	1367	2	I15766	[alpha-2A-adr
c 40	55.5	31.2	1372	2	I15767	[alpha-2A-adr
c 41	55.5	31.2	1374	2	I15768	[alpha-2A-adr
c 42	55.5	31.2	1461	2	I35766	[alpha-2A-adr
43	55	30.9	185	2	S55766	hypothetical prote
44	55	30.9	207	2	S34761	[alpha-2A-adr
45	55	30.9	474	2	I50766	hypothetical prote

ALIGNMENTS

RESULT 1
S12775
alpha-amylase (EC 3.2.1.1) precursor - [wheat, Triticum aestivum, cv. Sicco, germ] (fragment)
C;Species: Triticum aestivum (wheat)
C;Date: 30-Sep-1994 #sequence_revision 30-Sep-1994 #text_change 16-Jul-1999
C;Accession: S12775
R;O'Neill, S.D.; Kamaris, M.H.; Matthews, A.J.; Chandler, N.; Flavell, R.B.; Appels, R.
Mol. Gen. Genet. 221, 235-244, 1990
A;Title: The alpha amylase genes in cereal cultivars: characterization of a new Amy
A;Reference number: S12775; MUID:90370228; PMID:2374849
A;Accession: S12775
A;Molecule type: mRNA
A;Residues: 1-434 <ONE>
A;Cross-references: EMBL:M31405 ; NID:g166...; PIDN:AAA...; ...
C;Function:
A;Description: catalyzes the hydrolysis of internal 1,4 alpha-D-glucosidic bonds
A;Pathway: glycogen/starch degradation
C;Superfamily: wheat alpha-amylase; alpha-amylase
C;Keywords: glycosidase; hydrolase; polysaccharide degradation
F;1-29/Domain: alpha-amylase active site signature
F;209,244,430/Active site: Asp, Glu, Asp (catalytic nucleophile)

Alignment Scores:
Pred. No.: ... Length: ... Score: ... 146
Score: ... Matches: ... 91
Percent Similarity: ... Conservative: ...
Best Local Similarity: ... Mismatches: ...
Query Match: ... Indels: ... 0
DB: ... Gaps: ... 0

US-09-915-873-3 (1-96) x S12775 (1-434)

Qy 4 ATGCAGATCC...AACACTAGCC...ACGCGCGCGT 96
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 METQAVTLEGASNIMEMVA...SGLSGLSLEG...VA 434

Qy 97 TG 98
||
Db 435 TG 436

ed by the embryo during germination.
C;Genetics:
A;Gene: Amy2-2
A;Map position: 6
A;Introns: 29/3; 344/3
A;Note: multigene family on chromosome 6 encodes type B alpha-amylases; type A alpha-amy
C;Function:
A;Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A;Pathway: glycogen/starch degradation
C;Superfamily: wheat alpha-amylase; alpha-amylase core homology
C;Keywords: aleurone cell; blocked amino end; calcium binding; germination; glycosidase
F;1-24/Domain: signal sequence #status experimental <MA>
F;25-427/Product: alpha-amylase #status experimental <AMY>
F;170-316/Domain: alpha-amylase core homology <AMY>
F;25/Modified site: blocked amino end (Gln) (in mature form) (probably pyrrolidone carbo
F;203,228,313/Active site: Asp, Glu, Asp #status predicted

Alignment Scores
Pred. No.: 0.143 Length: 427
Score: 71.50 Matches: 16
Percent Similarity: 84.00% Conservative: 5
Best Local Similarity: 64.00% Mismatches: 3
Query Match: 40.17% Indels: 1
DB: 1 Gaps: 1

US-09-915-873-3 (1-96) x ALBHB (1-427)

Qy 22 ATCGTCAACAAGCACTTCCTCTCCCTGTCCGTCCTCATCGTGCTCCTCGCGCTGACCAGC 81
||| |||||||| |||||||||||||:: ::||||||||||||||||||||||:
Db 1 MetAlaAsnLysHis---LeuSerLeuSerLeuPheLeuValLeuLeuGlyLeuSerAla 19

Qy 82 AACCTCACCGCCGGC 96
:::||| :::|||
Db 20 SerLeuAlaSerGly 24

RESULT 5
JE0406
alpha-amylase (EC 3.2.1.1) B precursor (gene Amy56 and others) - barley
N;Alternate names: 1,4-alpha-D-glucan glucanohydrolase
C;Species: Hordeum vulgare (barley)
C;Date: 17-Apr-1993 #sequence_revision 21-Jan-1997 #text_change 18-Jun-1999
C;Accession: JE0406; B30759; S06275; B31960; B21826
R;Rahmatullah, R.J., Huang, J.K.; Clark, R.J.; Brady, C.K.; Chandra, G.K.; Muthukrishnan
Plant Mol. Biol. 12, 119-121, 1989
A;Title: Nucleotide and predicted amino acid sequences of two different genes for high-p
A;Reference number: JE0405
A;Accession: JE0406
A;Molecule type: DNA
A;Residues: 1-429 <RAH>
A;Cross-references: EMBL:X15227; NID:g18899; PIDN:CAA33211.1; [illegible]
A;Experimental source: gene Amy56 for alpha-amylase
A;Genetics: A56
R;Rogers, J.C.
submitted to GenBank, September 1988
A;Reference number: A94535
A;Accession: B30759
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-104,'G',106-154,'A',156-160,'PD',163,'G',167-196,'V',198-342,'I',344-392,'
A;Cross-references: GB:J04202; NID:g166984; PIDN:AAA32926.1; PID:g166985
A;Experimental source: cv. Himalaya gene Amy46 for alpha-amylase B
A;Genetics: A46
R;Knox, C.A.P.; Sonthayanon, B.; Chandra, G.K.; Muthukrishnan, S.
Plant Mol. Biol. 9, 3-17, 1987
A;Title: Structure and organization of two divergent alpha-amylase genes from barley.
A;Reference number: S06275
A;Accession: S06275
A;Molecule type: DNA
A;Residues: 1-11,'LI',14-32,'S',34-57,'I',59-79,'Y',81-139,'R',141-160,'PA',163-164,'R',
A;Cross-references: EMBL:M17125; NID:g166978; PIDN:AAA32926.1; PID:g166979
A;Experimental source: cv. Sundance gene for alpha amylase 1 precursor (clone pl41.117)
A;Genetics: AA1
A;Note: the authors translated the codon TCG for residue 33 as Trp
R;Khurshood, B.; Rogers, J.C.
J. Biol. Chem. 263, 18952-18960, 1988
A;Title: Barley alpha amylase genes: quantitative comparison of [illegible]
A;Reference number: A92700; MUID:89043007; PMID:3264263
A;Accession: B31960
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-51 <KHU>
A;Experimental source: strain Himalaya gene Amy46 for alpha amylase B [illegible]
R;Huang, J.K.; Swello, M.; Jacobsen, A.M.; Muthukrishnan, S.
J. Biol. Appl. Genet. 2, 579-588, 1984
A;Reference number: A92687; MUID:85110449; PMID:6335720
A;Accession: B21826
A;Molecule type: mRNA
A;Residues: 57-389,'I',391-392,'L',394-427 <HUA>
A;Cross-references: GB:J02636; NID:g166982; PIDN:AAA32926.1; PID:g166983
A;Experimental source: cv. Himalaya aleurone cell mRNA (clone [illegible] alpha amylase B)
C;Genetics: <A56>
A;Gene: Amy56
A;Introns: 29/3; 344/3
C;Genetics: <A46>
A;Gene: Amy46
C;Genetics: <AA1>
A;Gene: Amy1
A;Map position: 6
A;Introns: 29/3; 344/3
C;Function:
A;Description: catalyzes the hydrolysis of internal 1,4-alpha-[illegible]
A;Pathway: glycogen/starch degradation
C;Superfamily: wheat alpha-amylase; alpha-amylase core homology
C;Keywords: aleurone cell; germination; glycosidase; hydrolase; [illegible]
F;172-318/Domain: alpha-amylase core homology <AMY>
F;205,230,315/Active site: Asp, Glu, Asp #status predicted

Alignment Scores:
Pred. No.: 0.27 Length: 429
Score: 69.50 Matches: 15
Percent Similarity: 80.00% Conservative: 4
Best Local Similarity: 64.00% Mismatches: 3
Query Match: 38.61% Indels: 1
DB: 1 Gaps: 1

US-09-915-873-3 (1-96) x JE0406 (1-429)

Qy 22 ATCGTCAACAAGCACTTCCTCTCCCTGTCCGTCCTCATCGTGCTCCTCGCGCTGACCAGC 81
||| |||||||| :::||||||||||:: ||||||||||||||||||||||:
Db 1 MetAlaAsnLysHis---MetSerLeuSerLeuPheLeuValLeuLeuGlyLeuSerAla 19

Qy 82 AACCTCACCGCCGGC 96
:::||| :::|||
Db 20 SerLeuAlaSerGly 34

RESULT 6
T36169
hypothetical protein SCE2.09 - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change [illegible]
C;Accession: T36169
R;Seeger, K.J.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; [illegible]
Submitted to the EMBL Data Library, April 1999
A;Reference number: Z21599
A;Accession: T36169
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-486 <SEE>
A;Cross-references: EMBL:AL031064; PIDN:CAB11001.1; GSPDB:GN[illegible]
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SCOEDB:SCE2.09

Alignment Scores:
Pred. No.: 0.66 Length: 486
Score: 67.00 Matches: 15

A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues, 1-185 <TOM>
A;Cross-references: GB:AE000619; GB:AE000511; NID:g2511276; PIDN:AAD08168.1; PID:g2511427
C;Superfamily: PPPP-type peptidylprolyl isomerase homology <PPI>
F;10-57/Domain: BKBP-type peptidylprolyl isomerase homology <PPI>

Alignment Scores:
Pred. No.: 2.13
Score: 63.00
Percent Similarity: 54.17%
Best Local Similarity: 50.00%
Query Match: 37.28%
DB: 2
Length: 185
Matches: 12
Conservative: 1
Mismatches: 9
Indels: 2
Gaps: 1

US-09-915-873-3 (1-96) x C64660 (1-185)

Qy 85 GGTTGCTGCTCAGCCCGAGGACGACGATGACGACGACGGAGAGGAAGTGCTTGTTGA 26
Db 162 GlyCysCysGlyGlyHisGlyGlyHisGlyGlyLysLysGlyGlyGlyCysGlyCys-- 180

Qy 25 CCATCGTGTTCA 14
Db 181 ---SerCysSer 183

RESULT 11
T49385
hypothetical protein B1D1.390 [imported] - Neurospora crassa
C;Species: Neurospora crassa
C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
C;Accession: T49385
R;Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, May 2000
A;Reference number: Z25022
A;Accession: T49385
A;Status: preliminary
A;Molecule type: DNA
A;Residues, 1-805 <SCH>
A;Cross-references: EMBL:AL355927; GSPDB:GN00110; NID:g3881200
A;Experimental source: BAC clone B1D1; strain OR74A
C;Genetics:
A;Gene: NCSP:B1D1.390
A;Map position: 6
A;Introns: 54/3; 212/3

Alignment Scores:
Pred. No.: 2.55
Score: 62.50
Percent Similarity: 50.00%
Best Local Similarity: 50.00%
Query Match: 35.11%
DB: 2
Length: 805
Matches: 15
Conservative: 6
Mismatches: 8
Indels: 7
Gaps: 2

US-09-915-873-3 (1-96) x T49385 (1-805)

Qy 12 CCTGAACACGATGGTCAACAAGCACTTCCTCTCCCTGTCCGTCCTCATCGT--------- 62
Db 198 ProProHisGlyGlyAsnThrAlaPheProLeu------ArgProHisAsnGlyProAsp 215

Qy 63 ------CCTCCTCGGGCTGAGCAGCAACCT 86
Db 216 ValLeuProProArgAlaThrThrGlnPro 225

RESULT 12
T32594
hypothetical protein C02B10.5 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C;Accession: T32594
R;Nelson, J.; White, S.; Hawkins, J.; Wohldmann, P.
submitted to the EMBL Data Library, December 1997
A;Description: The sequence of C. elegans cosmid C02B10.
A;Reference number: Z21196
A;Accession: T32594
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-698 <NEL>
A;Cross-references: EMBL:AF039805; PIDN:AAB94202.1; GSPDB:GN00110; NID:g3881200
A;Experimental source: strain Bristol N2; clone C02B10
C;Genetics:
A;Gene: CESP:C02B10.5
A;Map position: 4
A;Introns: 61/3; 102/2; 168/3; 449/2; 541/3

Alignment Scores:
Pred. No.: 3.5
Score: 61.50
Percent Similarity: 41.46%
Best Local Similarity: 34.15%
Query Match: 34.56%
DB: 2
Length: 698
Matches: 14
Conservative: 5
Mismatches: 11
Indels: 14
Gaps: 2

US-09-915-873-3 (1-96) x T32594 (1-698)

Qy 4 CATGCAGGT- ---- ---- ----AAACAGATGGTCAACAAGCACTTCCT 34
Db 602 HisProGlyMetAlaGlyMetProGlyMetHisProGlyMetGlnGlyMetPro-Pro 640

Qy 42 ---- ---- ---- ---CTCCCTCTCCCTGTCCGTCCTCATCGTCCTCCTCGGGCTGAGCAGCAACCT 86
Db 622 HisProGlyMetGlnGlyMetProGlyMetProHisGlyProProMetProPro-ProAsn 681

Qy 84 CCT 86
Db 642 Pro 642

RESULT 13
I69193
ataxin 7 - human
C;Species: Homo sapiens (man)
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 25-Jun-1999
C;Accession: I69193
R;Del-Favero, J.; Krols, L.; Michalik, A.; Theuns, J.; Lubotzki, A.; Goossens, V. Web
.
Hum. Mol. Genet. 7, 177-186, 1998
A;Title: Molecular genetic analysis of autosomal dominant cerebellar ataxia type
A;Reference number: Z16604; MUID:98208504; PMID:9425224
A;Accession: I69193
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-892 <DEL>
A;Cross-references: EMBL:AF032185; NID:g3192953; PIDN:AAC39956.1; PID:g3192954
A;Experimental source: cell type lymphoblast
C;Genetics:
A;Gene: SCA7
A;Map position: 3

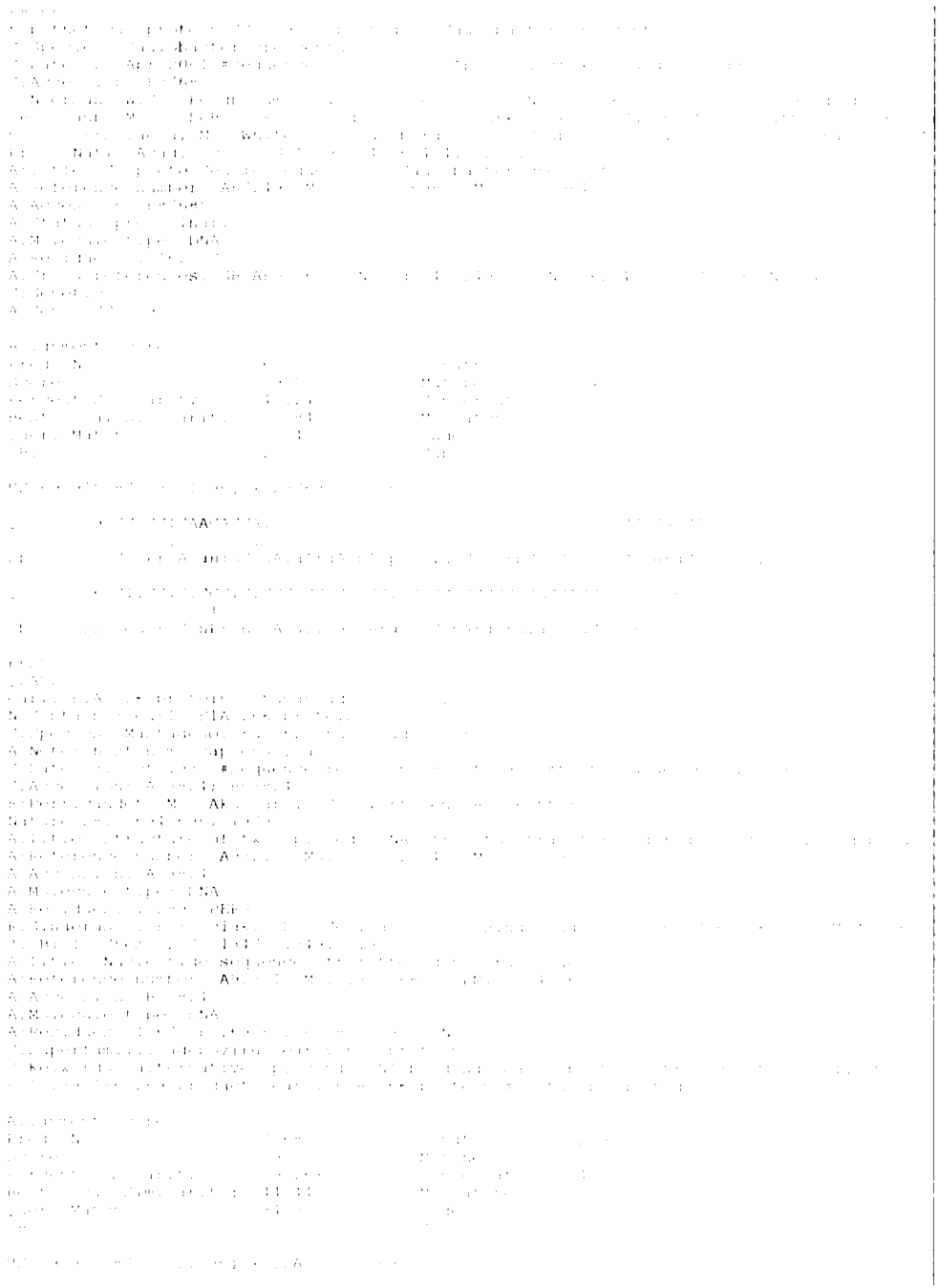
Alignment Scores:
Pred. No.: 5.6
Score: 60.00
Percent Similarity: 59.26%
Best Local Similarity: 48.15%
Query Match: 34.71%
DB: 2
Length: 892
Matches: 13
Conservative: 3
Mismatches: 9
Indels: 2
Gaps: 2

US-09-915-873-3 (1-96) x I69193 (1-892)

Qy 18 CACGATGGTCAACAAGCACTTCCTCTCCCTGTCCGTCCTCATC 61
Db 401 HisProAspSerGlnGlnProLeuAlaProLeuArgAspProHisPro-AlaProLeuArg 459

Qy 72 GCTGAGCAGCAACCTCACCGC 92
Db 421 ThrSerGlnGlnProHisGlnGly 442

RESULT 14



GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: January 17, 2003, 05:22:27 ; Search time 52.192 Seconds
(without alignments)
479.974 Million cell updates/sec

Title: US-09-915-873-4
Perfect score: 960
Sequence: 1 MALTFALLVALLVLSCKSSC.....EIMNSFSLSTNLQESLRSKE 188

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseq-aa.mlf:* *
2: /SIDS2/gcgdata/geneseq/geneseq-aa.mlf:* *
3: /SIDS2/gcgdata/geneseq/geneseq-aa.mlf:* *
4: /SIDS2/gcgdata/geneseq/geneseq-aa.mlf:* *
5: /SIDS2/gcgdata/geneseq/geneseq-aa.mlf:* *
6: /SIDS2/gcgdata/geneseq/geneseq-aa.mlf:* *
7: /SIDS2/gcgdata/geneseq/geneseq-aa.mlf:* *
8: /SIDS2/gcgdata/geneseq/geneseq-aa.mlf:* *
9: /SIDS2/gcgdata/geneseq/geneseq-aa.mlf:* *
10: /SIDS2/gcgdata/geneseq/geneseq-aa.mlf:* *
11: /SIDS2/gcgdata/geneseq/geneseq-aa.mlf:* *
12: /SIDS2/gcgdata/geneseq/geneseq-aa.mlf:* *
13: /SIDS2/gcgdata/geneseq/geneseq-aa.mlf:* *
14: /SIDS2/gcgdata/geneseq/geneseq-aa.mlf:* *
15: /SIDS2/gcgdata/geneseq/geneseq-aa.mlf:* *
16: /SIDS2/gcgdata/geneseq/geneseq-aa.mlf:* *
17: /SIDS2/gcgdata/geneseq/geneseq-aa.mlf:* *
18: /SIDS2/gcgdata/geneseq/geneseq-aa.mlf:* *
19: /SIDS2/gcgdata/geneseq/geneseq-aa.mlf:* *
20: /SIDS2/gcgdata/geneseq/geneseq-aa.mlf:* *
21: /SIDS2/gcgdata/geneseq/geneseq-aa.mlf:* *
22: /SIDS2/gcgdata/geneseq/geneseq-aa.mlf:* *
23: /SIDS2/gcgdata/geneseq/geneseq-aa.mlf:* *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

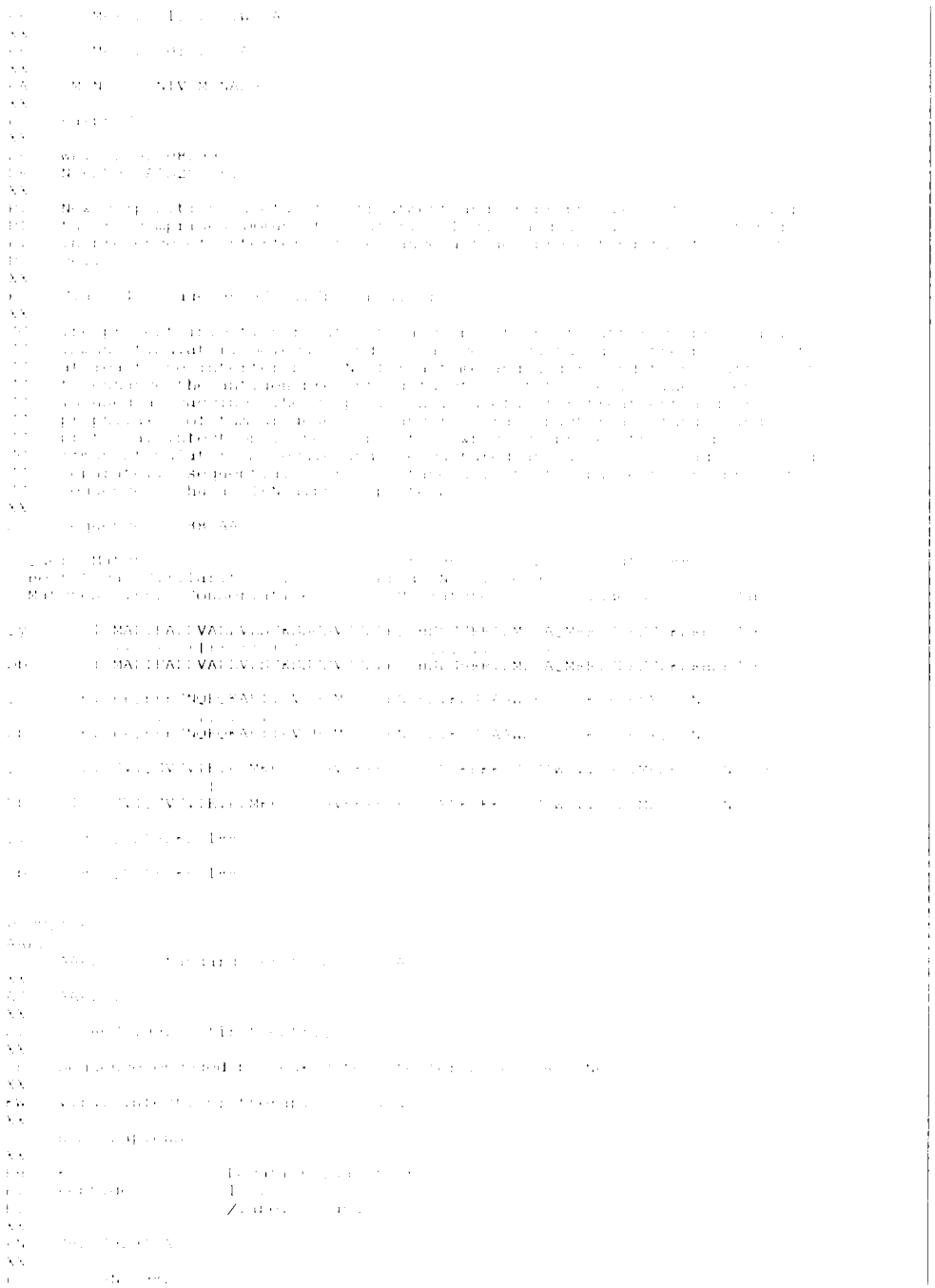
SUMMARIES

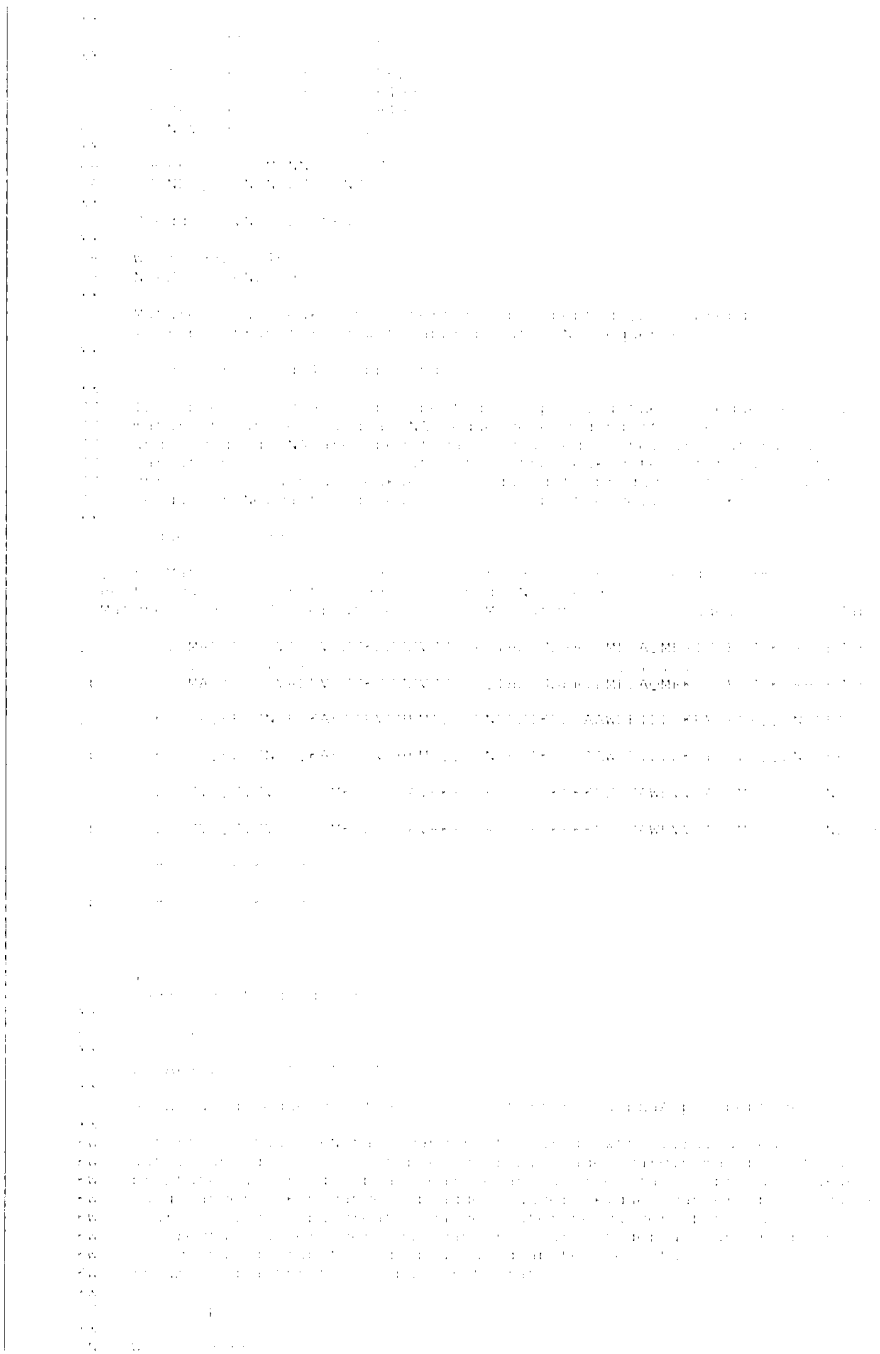
Result No.	Score	Query Match	Length	DB	ID	Description
1	960	100.0	188	20	AAY42485	Human interferon p
2	960	100.0	188	21	AAY69484	Amino acid sequen
3	960	100.0	188	23	AAE18957	Human alpha-2b-int
4	960	100.0	188	23	ABB07434	Interferon alpha2
5	960	100.0	188	23	AAE15828	Human interferon (
6	957	99.7	189	3	AAP20103	Sequence encoded b
7	957	99.7	189	22	AAB21458	Amino acid sequenc
8	957	99.7	188	23	ABB80767	Human interferon a
9	957	99.7	189	23	AAE15826	Human interferon (
10	957	99.7	190	22	AAB82438	Human interferon a
11	957	99.7	216	4	AAP08515	New type I interf
12	954	99.4	188	20	AAE19695	Interferon alpha
13	952	99.2	188	23	AAP20514	Protein sequence
14	952	99.2	188	23	AAP20516	Protein sequence N
15	944	98.3	187	4	AAP20007	Hybrid interferon
16	943	98.2	188	22	AAB21455	Amino acid sequenc
17	942	98.1	190	2	AAP10018	Sequence 3 from Pa
18	939	97.8	175	1	AAL13737	Sequence 3 from Pa
19	867	90.3	344	23	ABB00414	Human interferon a
20	855.5	89.1	527	22	AAB21457	Unit 1 of fusion N
21	853	88.9	271	20	AAY42486	Amino acid sequen
22	853	88.9	455	2	AAY69485	Amino acid sequen
23	852	88.8	527	21	AAY69483	Unit 1 of fusion
24	852	88.8	527	23	AAE18956	Human alpha-2b-int
25	852	88.8	271	23	AAE15827	Human interferon (
26	851.5	88.7	527	19	AAP03533	Interferon alpha
27	851	88.6	344	4	AAP20105	Interferon alpha
28	851	88.6	344	19	AAP03534	Recombinant human
29	851	88.6	344	20	AAB21459	Human interferon
30	851	88.6	344	20	AAY42487	Interferon alpha
31	851	88.6	344	21	AAY69486	Amino acid sequen
32	851	88.6	344	22	AAY19494	Amino acid sequen
33	851	88.6	344	23	AAY19493	Amino acid sequen
34	851	88.6	344	23	AAY 2	Amino acid sequen
35	851	88.6	345	20	AAY42488	Amino acid sequen
36	851	88.6	345	21	AAY69487	Interferon alpha
37	851	88.6	345	21	AAY69488	Interferon alpha
38	851	88.6	165	21	AAY69489	N-terminal 95 AA
39	851	88.6	165	23	AAE18955	Human alpha-2b-int
40	851	88.6	165	23	AAM50705	Human alpha-2b-int
41	85	88.6	165	23	AAM50706	Human interferon a
42	851	88.6	345		AAW22134	Human interferon a
43	851	88.6	345		AAP07507	Sequence 3 from Pa
44	851	88.6	344	19	AAW22133	Protein #3 - Unkno
45	851	88.6	345		AAW22135	Protein #3 - Unkno

ALIGNMENTS

RESULT 1
AAY42485
standard; protein; 188 AA.
XX
AC AAY42485;
XX
DT 25-FEB-2000 (first entry)
XX
DE Human interferon protein variant.
XX
KW WHO-type human interferon; alpha interferon; cytokine; [illegible]
KW pharmaceutical composition; [illegible] disease; [illegible]
KW cancer.
XX
OS Homo sapiens.
XX
PN WO9947678-A2.
XX
PD 23-SEP-1999.
XX
PF 12-MAR-1999; 99WO-US005411.
XX
PR 19-MAR-1998; 98US-078441.
XX
PA (GENE) GENEMEDICINE INC.
XX
PI Nordstrom J, Tetrault P, [illegible] A, Silvestro P;
XX
DR WPI; 1999-562117/47.
DR N-PSDB; AAY40412.
XX
PT New plasmids containing interferon alpha subtype sequences, useful [illegible]
PT the treatment of a cancer in a [illegible] disease [illegible]

```
XX
DR   WPI; 2002-195966/25.
XX
PT   Producing recombinant polypeptides from duckweed plant culture, by
PT   transforming culture with nucleotide sequence coding for the
PT   polypeptide and signal peptide that directs polypeptide secretion into
PT   culture medium -
XX
PS   Claim 21; Page 45-46; 47pp; English.
XX
CC   The invention relates to a method for producing a biologically active
CC   recombinant polypeptide. The method comprises culturing a duckweed plant
CC   culture or duckweed nodule culture, which is stably transformed to
CC   express the polypeptide encoded by a nucleotide sequence that has been
CC   modified for enhanced expression in duckweed and collecting the
CC   polypeptide from duckweed plant or nodule culture. The method is useful
CC   for producing a biologically active recombinant polypeptide and for the
CC   directed secretion of the polypeptide from duckweed plant or nodule
CC   cultures. The polypeptides include a mammalian, therapeutic polypeptide
CC   such as insulin, interferons, monoclonal antibodies, cytokines, vaccines,
CC   in particular human alpha-2b-interferon or its biologically active
CC   variant. The duckweed frond culture or duckweed nodule culture expresses
CC   and assembles all of the subunits of a biologically active multimeric
CC   protein chosen from collagen, haemoglobin, P450 oxidase, and a monoclonal
CC   antibody, useful for industrial or chemical processes or for diagnostic,
CC   therapeutic or vaccination purposes. Nucleic acid encoding the
CC   recombinant polypeptides are useful for the expression and secretion of
CC   human alpha-2b-interferon in duckweed. The present sequence is human
CC   alpha-2b-interferon precursor protein.
XX
SQ   Sequence   188 AA;

  Query Match             100.0%;  Score 960;  DB 23;  Length 188;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-90;
  Matches  188;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MALTFALLVALLVLSCKSSCSVGCDLPQTHSLGSRRTLMLLAQMR[illegible]
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MALTFALLVALLVLSCKSSCSVGCDLPQTHSLGSRRTLMLLAQMR[illegible]

Qy         61 FPQEEFGNQFQKAETIPVLHEMIQQIFNLFSTKDSSAAWDETLLDKFYTELYQQLN[illegible]
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 FPQEEFGNQFQKAETIPVLHEMIQQIFNLFSTKDSSAAWDETLLDKFYTELYQQLN[illegible]

Qy        121 CVIQGVGVTETPLMKEDSILAVRKYFQRITLYLKEKKYS[illegible]
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CVIQGVGVTETPLMKEDSILAVRKYFQRITLYLKEKKYS[illegible]

Qy        181 QESLRSKE 188
              ||||||||
Db        181 QESLRSKE 188


RESULT 4
ABB07434
ID   ABB07434 standard; peptide; 188 AA.
XX
AC   ABB07434;
XX
DT   09-APR-2002  (first entry)
XX
DE   Interferon-alpha2 protein fragment.
XX
KW   Interferon-beta-2; IFN-beta2; neuroprotective; cytostatic; virucide;
KW   antiarthritic; antirheumatic; gene therapy; interferon-alpha2.
XX
OS   Unidentified.
XX
PN   WO200195929-A2.
XX
PD   20-DEC-2001.
XX
PF   18-JUN-2001; 2001WO-US41022.
XX
PR   16-JUN-2000; 2000US-[illegible]
PR   15-JUN-2001; [illegible]
XX
PA   (SCHD ) SCHERING AG.
XX
PI   [illegible]
XX
DR   WPI; [illegible]
XX
PT   Composition for treating multiple sclerosis, cancer and [illegible]
PT   and infections, comprises human interferon-beta-2 or its
PT   biologically-active fragment or [illegible]
XX
PS   Disclosure; Fig 4; 61pp; English.
XX
CC   The invention relates to a pharmaceutical composition [illegible]
CC   therapeutically effective amount of human interferon-beta-2 (IFN-beta2)
CC   polypeptide. The composition is useful for treating multiple sclerosis in
CC   mammals, in particular a human in need of such treatment, and also cancer
CC   e.g. intraepithelial neoplasia and cervical cancer, autoimmune diseases
CC   e.g. rheumatoid arthritis and viral diseases or infections. The
CC   composition is useful for anti-oncogene regulation, antitumour activity,
CC   antiviral activity, cell growth inhibition or antigrowth activity, anti-
CC   proliferation, enhancement of cytotoxicity of lymphocytes, [illegible]
CC   inhibition of differentiation of target cells, immunoregulatory activity,
CC   macrophage activation and down regulation of oncogenes. [illegible]
CC   ABB07427-441 represent various interferon (IFN) sequences used for
CC   alignment studies with the human IFN-beta2 polypeptide.
XX
SQ   Sequence   188 AA;

  Query Match             [illegible]
  Best Local Similarity   [illegible]
  Matches  188;  Conservative  [illegible]

Qy          1 [illegible]
Db          1 [illegible]

Qy         61 [illegible]
Db         61 [illegible]

Qy        121 [illegible]
Db        121 [illegible]

Qy        181 QESLRSKE 188
              ||||||||
Db        181 QESLRSKE 188


RESULT 5
AAE15828
ID   AAE15828 standard; protein; 188 AA.
XX
AC   AAE15828;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Human interferon (IFN) [illegible]
XX
KW   Human; vaccine; [illegible]
KW   antigen presentation; [illegible]
KW   antitumour; antibacterial; [illegible]
XX
OS   Homo sapiens.
XX
PN   WO200189097-A1.
XX
PD   22-NOV-2001.
XX
```





XX
PD 28-DEC-2000.
XX
PF 22-JUN-2000; 2000WO-IB01000.
XX
PR 22-JUN-1999; 99US-0140411
XX
PA (UYMA-) UNIV MARYLAND BALTIMORE.
XX
PI Pontzer CH, Shorts LH, Clark CD;
XX
DR WPI; 2001-071357/08
XX
PT Producing recombinant interferon tau analog proteins with improved
PT properties, useful for treating cancers, autoimmune diseases and viral
PT infections -
XX
PS Claim 6; Page 60-61; 70pp; English.
XX
CC The present sequence represents a human interferon alphaA polypeptide.
CC The specification describes a method of making recombinant interferon
CC (IFN)-tau proteins, which differ from wild-type IFN-tau by amino acid
CC substitutions near the amino terminus of the molecule. The mutated
CC IFN-tau proteins have improved biological activity, low toxicity,
CC retain the same or slightly reduced antiviral activity compared with
CC interferon alpha, and have enhanced antiproliferative activity compared
CC to wild-type IFN-tau. The method is used for producing IFN-tau proteins
CC with improved biological activities and properties. These IFN-tau may
CC be administered to treat cancers and decrease tumour growth, treat
CC viral diseases, treat autoimmune diseases and treat multiple sclerosis.
CC The cancer or tumour is selected from the group comprising human
CC adenocarcinoma, breast cancer, prostate cancer, glioblastomas, melanomas,
CC myelomas, lymphomas, leukaemia, lung cancer, skin cancer, bladder cancer,
CC kidney cancer, brain cancer, ovarian cancer, pancreatic cancer, uterine
CC cancer, bone cancer, colorectal cancer, cervical cancer and
CC neuroectodermal cancer, monoclonal gammapathies and cervical and oral
CC dysplasia. The autoimmune disease is selected from Type I diabetes
CC mellitus, rheumatoid arthritis, lupus erythematosus and/or psoriasis.
CC The viral infection is an RNA virus, a human immunodeficiency virus
CC (HIV) or hepatitis C virus.
XX
SQ Sequence 188 AA;

Query Match 99.7%; Score 957; DB 22; Length 188;
Best Local Similarity 99.5%; Pred. No. 5e-98;
Matches 187; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MALTFALLVALLVLSCKSSCSVGCDLPQTHSLGSRRTLMLLAQMRRISLFSCLKDRHDFG 60
Db 1 MALTFALLVALLVLSCKSSCSVGCDLPQTHSLGSRRTLMLLAQMRKISLFSCLKDRHDFG 60

Qy 61 FPQEEFGNQFQKAETIPVLHEMIQQIFNLFSTKDSSAAWDETLLDKFYTELYQQLNDLEA 120
Db 61 FPQEEFGNQFQKAETIPVLHEMIQQIFNLFSTKDSSAAWDETLLDKFYTELYQQLNDLEA 120

Qy 121 CVIQGVGVTETPLMKEDSILAVRKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTNL 180
Db 121 CVIQGVGVTETPLMKEDSILAVRKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTNL 180

Qy 181 QESLRSKE 188
Db 181 QESLRSKE 188

RESULT 8
ABB80767
ID ABB80767 standard; Protein; 188 AA.
XX
AC ABB80767;
XX
DT 23-SEP-2002 (first entry)
XX
DE Human interferon alpha 2 protein.
XX
KW Human; interferon; ,provides; [illegible] activity; [illegible]
KW antiproliferative protein; antiviral; interferon alpha 2.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 36
FT /note= "Cys-Ala-Asp; A-A"
XX
PN WO200036015-A1.
XX
PD 15-MAY-2000.
XX
PF 26-SEP-2000; 2000WO-JP06473.
XX
PR 27-OCT-1999; 99JP-0305295.
PR 16-APR-2000; 2000JP-0114775.
XX
PA (HAYA-) HAYASHIBARA SEIBUTSU KAGAKU.
XX
PI Yamai Y, Sano T, Kudo Y, Mori M;
XX
DR WPI; 2000-375275/32.
DR N-PSDB; ABN90168.
XX
PT Expression potentiators for protein synthesis regulatory genes -
PT medicines containing subtypes alpha2 and alpha8 of human interferon
PT alpha, for treating, ameliorating or preventing diseases of brain, nerves
PT and circulatory systems -
XX
PS Disclosure; Page 11; 24pp; Japanese.
XX
CC The invention relates to an expression potentiator for a protein
CC synthesis-regulatory gene containing subtypes alpha2 and alpha8 of human
CC interferon alpha as active ingredients. The expression potentiator is
CC for protein synthesis regulatory genes containing subtypes alpha2 and
CC alpha8 of human interferon alpha, and are useful for treating, preventing
CC or ameliorating diseases of brain, digestive, nervous and circulatory
CC systems, skin, hematopoietic tissues and muscles, virus infections,
CC arteritis, Acquired immune deficiency syndrome, allergy, and cancers.
CC Medicines containing the peptides are stable, easily handled, and
CC highly active. The present sequence represents a human interferon alpha
CC alpha 2 protein.
XX
SQ Sequence 188 AA;

Query Match 99.7%; Score 957; DB 23; Length 188;
Best Local Similarity 99.5%; Pred. No. 5e-98;
Matches 187; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

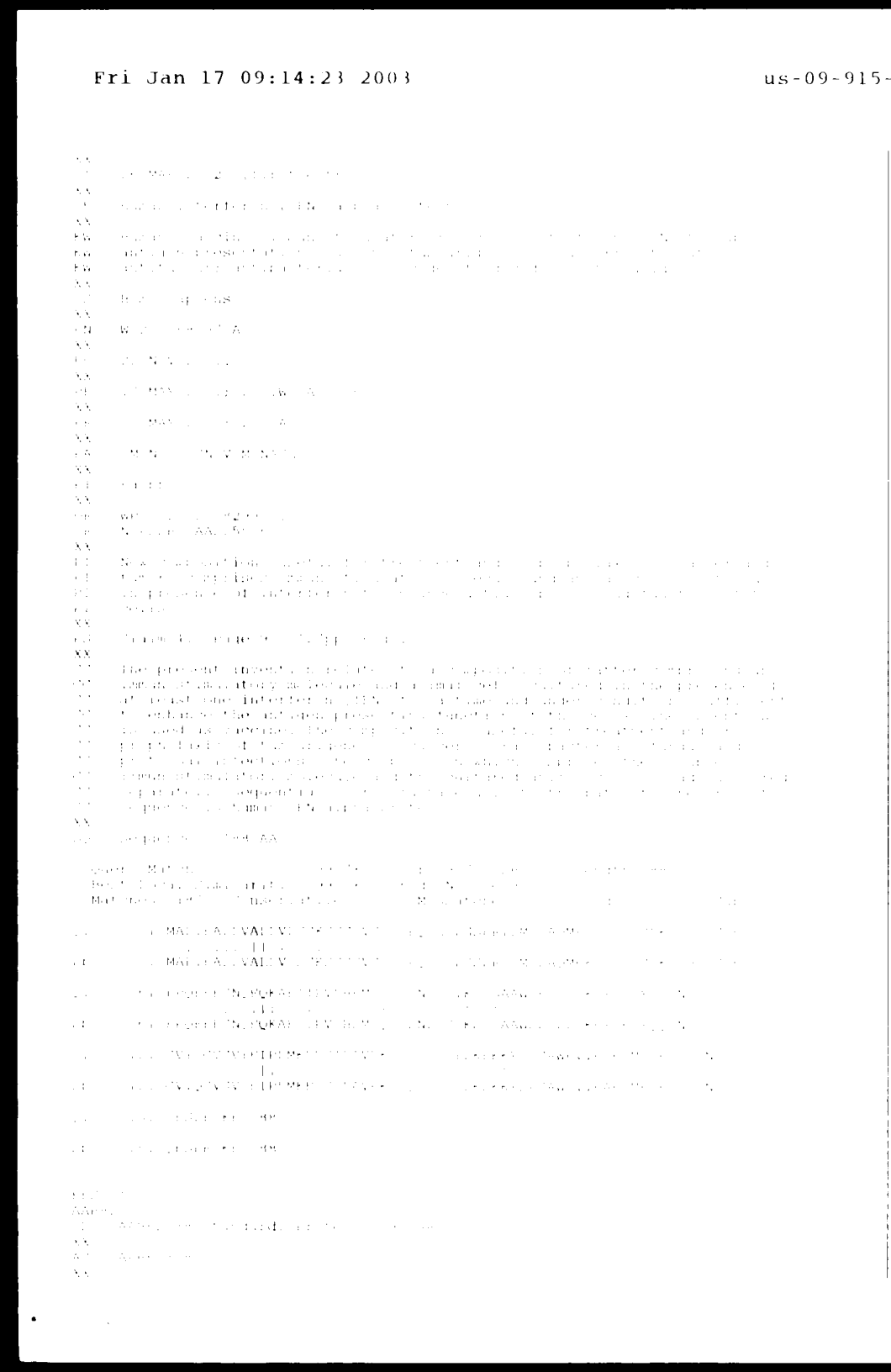
Qy 1 MALTFALLVALLVLSCKSSCSVGCDLPQTHSLGSRRTLMLLAQMRRISLFSCLKDRHDFG 60
Db 1 MALTFALLVALLVLSCKSSCSVGCDLPQTHSLGSRRTLMLLAQMRKISLFSCLKDRHDFG 60

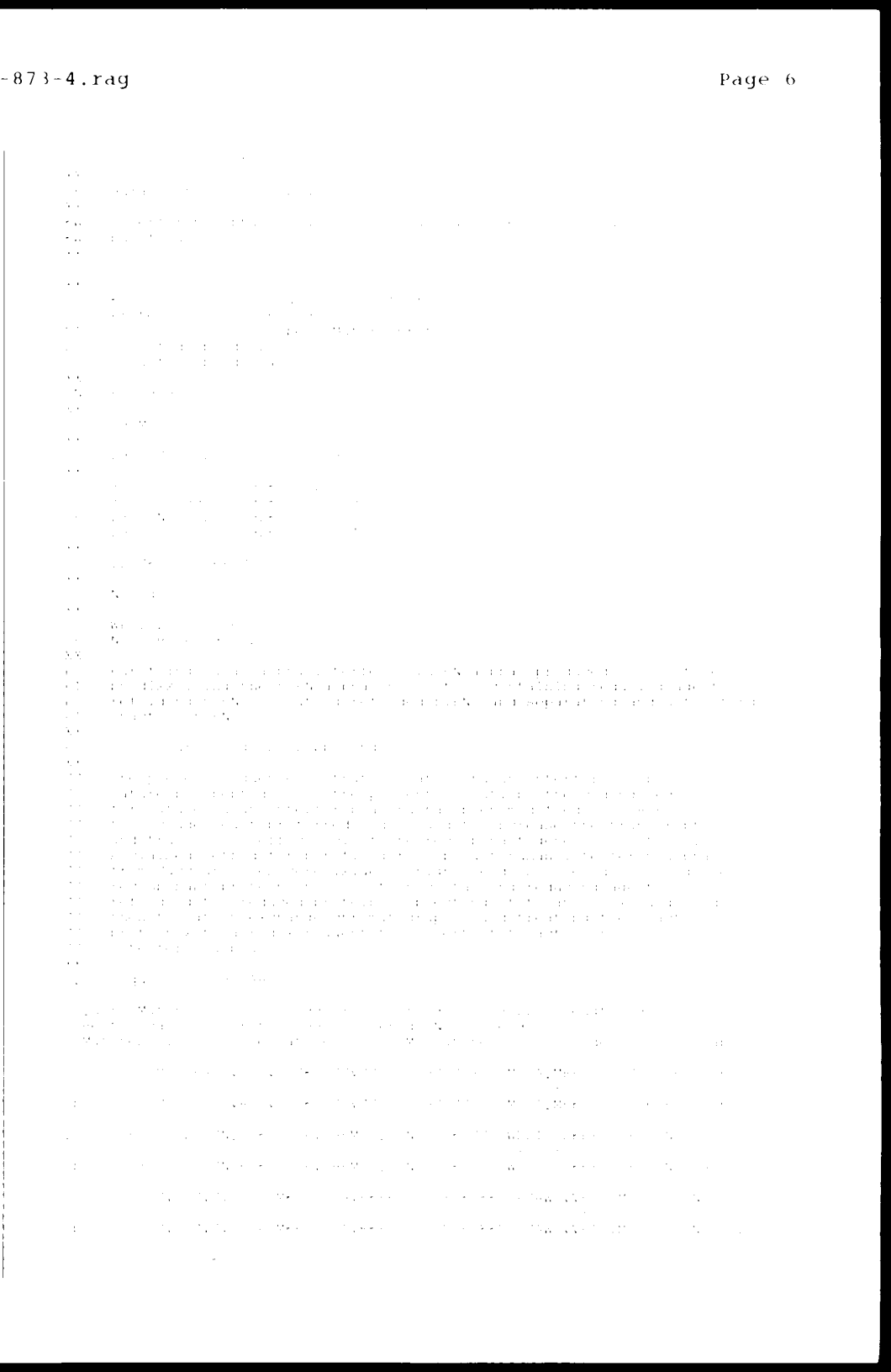
Qy 61 FPQEEFGNQFQKAETIPVLHEMIQQIFNLFSTKDSSAAWDETLLDKFYTELYQQLNDLEA 120
Db 61 FPQEEFGNQFQKAETIPVLHEMIQQIFNLFSTKDSSAAWDETLLDKFYTELYQQLNDLEA 120

Qy 121 CVIQGVGVTETPLMKEDSILAVRKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTNL 180
Db 121 CVIQGVGVTETPLMKEDSILAVRKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTNL 180

Qy 181 QESLRSKE 188
Db 181 QESLRSKE 188

RESULT 9
AAB15826
ID AAB15826 standard; Protein; 188 AA.
XX
AC AAB15826;





```
Db    183 QESLRSKE 190


RESULT 11
AAR38793
ID   AAR38793 standard; Protein; 216 AA.
XX
AC   AAR38793;
XX
DT   21-JAN-1994  (first entry)
XX
DE   Natural human interferon-alpha.
XX
KW   IFN-alpha; hIFN-alpha.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Region          26
FT                   /note= "Natural mature human alpha-interferon N-terminal"
FT   Region          190
FT                   /note= "Natural mature human alpha-interferon C-terminal"
FT   Region          191
FT                   /note= "corresponds to termination codon"
FT   Region          211
FT                   /note= "corresponds to termination codon"
XX
PN   EP553494-A.
XX
PD   04 AUG-1993.
XX
PF   29-DEC-1992;   92EP-0122084.
XX
PR   31-DEC-1991;   91KR-0025878.
PR   31-DEC-1991;   91KR-0025879.
PR   28-JAN-1992;   92KR-0001155.
XX
PA   (LUCK-) LUCKY LTD.
XX
PI   Bae TO,  Chang HJ,  Cho JM,  Park SJ,  Park YW.
XX
DR   WPI; 1993-244548/31.
DR   N-PSDB; AAQ47153.
XX
PT   Recombinant human alpha interferon and corresp. gene - for
PT   efficient expression in yeast, and purificn. process
XX
PS   Disclosure; Fig 1B; 30pp; English.
XX
CC   The sequence is that of natural human interferon-alpha.
XX
SQ   Sequence   216 AA;

  Query Match             99.7%;  Score 957;  DB 14;  Length 216;
  Best Local Similarity   99.5%;  Pred. No. 3.6e-90;
  Matches  187;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 MAITPALLVALLVLSCKSSCSVGCDLPQTHSLGSRRTLMLLAQMRRISPFSCLKDRHDFG 60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3 MAITPALLVALLVLSCKSSCSVGCDLPQTHSLGSRRTLMLLAQMRRISPFSCLKDRHDFG 62

Qy       61 FPQEEFGNQFQKAETIPVLHEMIQQIFNLFSTKDSSAAWDETLLDKFYTELYQQLNDLEA 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       63 FPQEEFGNQFQKAETIPVLHEMIQQIFNLFSTKDSSAAWDETLLDKFYTELYQQLNDLEA 122

Qy      121 CVIQGVGVTETPLMKEDSILAVRKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSINL 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      123 CVIQGVGVTETPLMKEDSILAVRKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSINL 182

Qy      181 QESLRSKE 188
            ||||||||
Db      183 QESLRSKE 190


RESULT 12
AAW18579
ID   AAW18579 standard; Protein; 193 AA.
XX
AC   AAW18579;
XX
DT   24-DEC-1997  (first entry)
XX
DE   Interferon alpha immunogen with 3 mutations.
XX
KW   Linker peptide; fusion protein; interferon alpha immunogen; IL-1;
KW   PCR; cytokine; peptide; recombinant DNA; T cell; vaccine;
KW   multiple myeloma; cancer; viral disease; virology.
XX
OS   Chimeric Homo sapiens.
OS   Chimeric synthetic.
XX
FH   Key             Location/Qualifiers
FT   Protein         1..193
FT                   /label= IFN-alpha
FT   Peptide         193..204
FT                   /label= linker
FT                   /note= "synthetic"
FT   Protein         205..415
FT                   /label= IL-1
XX
PN   WO9749417-A1.
XX
PD   31-DEC-1997.
XX
PF   24-DEC-1997;   97WO-US023001.
XX
PR   25-JUN-1996;   96US-0020211.
PR   26-SEP-1996;   96US-0026911.
XX
PA   (IMMU-) IMMUNEX RES DEVELOPMENT CORP.
XX
PI   Bauer JW,  Yu L;
XX
DR   WPI; 1998-064572/06.
DR   N-PSDB; AAV62924.
XX
PT   Hybrid of interferon and interleukin-1 for linked use
PT   non-immunogenic peptide - to treat e.g. viral infection
PT   or cancer, by fusing cell cytokine and IL-1 gene
XX
PS   Claim 1; Page 75-76; 92pp; English.
XX
CC   This polypeptide sequence comprises a fusion protein of
CC   human interferon-alpha (IFN-alpha) joined at its C-terminal end
CC   via a claimed linker peptide (see AAW18578) to the N-terminal end
CC   of human immunofobrin heavy chain gamma 1 (IC). It is prepared by
CC   PCR amplification (see AAW18581-21) of interferon alpha and IC
CC   sequences, construction of an expression vector, and expression in
CC   mammalian cells. The linker is a T cell immunologically inert
CC   peptide that serves to eliminate the non-antigenicity, protects
CC   the junction of the two polypeptide moieties. It also keeps
CC   coverage steric hindrance that the IC portion of the hybrid.
CC   hybrid has a much longer in vivo half-life than native IFN-
CC   alpha and will not penetrate membrane sites. It can be used
CC   treat hepatitis B + C, solid cell carcinoma, multiple myeloma
CC   or other cancers or viral diseases (claimed). It also serves as
CC   a model for the design and construction of other cytokine-IL-1
CC   hybrids.
XX
SQ   Sequence   415 AA;

  Query Match             49.4%;  Score 474;  DB 14;  Length 415;
  Best Local Similarity   49.0%;  Pred. No. 1.2e-40;
  Matches  187;  Conservative   21;  Mismatches   76;  Indels    3;  Gaps    1;

Qy        1 MAITPALLVALLVLSCKSSCSVGCDLPQTHSLGSRRTLMLLAQMRRISPFSCLKDRHDFG 60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3 MAITPALLVALLVLSCKSSCSVGCDLPQTHSLGSRRTLMLLAQMRRISPFSCLKDRHDFG 62

Qy       61 FPQEEFGNQFQKAETIPVLHEMIQQIFNLFSTKDSSAAWDETLLDKFYTELYQQLNDLEA 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       63 FPQEEFGNQFQKAETIPVLHEMIQQIFNLFSTKDSSAAWDETLLDKFYTELYQQLNDLEA 122

Qy      121 CVIQGVGVTETPLMKEDSILAVRKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSINL 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      123 CVIQGVGVTETPLMKEDSILAVRKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSINL 182

Qy      181 QESLRSKE 188
            ||||||||
Db      183 QESLRSKE 190
```

```
Db     121 QVIQGVGVTETPLMKEDSILAVRKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTN  180

Qy     181 QESLRSKE 188
           ||||||||
Db     181 QESLRSKE 188


RESULT 15
AAP20007
ID   AAP20007 standard; Protein; 187 AA.
XX
AC   AAP20007;
XX
DT   18-DEC-1992  (first entry)
XX
DE   Hybrid human leukocyte interferon LoIFA.
XX
KW   Leukocyte; interferon; antitumor; immunostimulant; virucide; plasmid;
KW   pLe-IFA.
XX
OS   Homo sapiens.
XX
PN   EP51873-A.
XX
PD   19-MAY-1982.
XX
PF   09-NOV-1981;   81EP-0109579.
XX
PR   25-SEP-1981;   81US-0305657.
PR   10-NOV-1980;   80US-0205579.
PR   23-FEB-1981;   81US-0237488.
XX
PA   (GENE-) GENENTECH INC.
XX
PI   Goeddel DVN;
XX
DR   WPI; 1982-41788E/21  (41788E).
DR   N-PSDB; AAN20005.
XX
PT   Hybrid human leukocyte interferon(s) - useful for treating viral
PT   and neoplastic diseases
XX
PS   Disclosure; Fig 1; 54pp; English.
XX
CC   This protein is expressed in Escherichia coli using the replicable
CC   expression vector plasmid pLe-IFA. See also AAN20006-13, AAN20020-[illegible]
CC   and AAP20008-14.
XX
SQ   Sequence   187 AA;

  Query Match             98.38%;  Score 944;  DB 5;  Length 187;
  Best Local Similarity   98.46%;  Pred. No. 6.5e-64;
  Matches  184;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 ALIFALLVALLVLSCKSSCSVGCDLPQTHSLGSRRTLMLLAQMRRISLFSCLKDRHDFGF 61
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ALIFALLVALLVLSCKSSCSVGCDLPQTHSLGSRRTLMLLAQMRRISLFSCLKDRHDFGF 60

Qy      62 PQEEFGNQFQKAETIPVLHEMIQQIFNLFSTKDSSAAWDETLLDKFYTELYQQLNDLEAC 121
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 PQEEFGNQFQKAETIPVLHEMIQQIFNLFSTKDSSAAWDETLLDKFYTELYQQLNDLEAC 120

Qy     122 VIQGVGVTETPLMKEDSILAVRKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTNLQ 181
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 VIQGVGVTETPLMKEDSILAVRKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTNLQ 180

Qy     182 ESLRSKE 188
           ||||| |
Db     181 ESLKSNE 187


Search completed: January 17, 2003, 06:53:54
Job time : 59.1926 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2003, 05:09:36 ; Search time 27.00 Seconds
(without alignments)
5892.893 Million cell updates/sec

Title: US-09-915-873-1
Perfect score: 554
Sequence: 1 gatcaaqtacaaaaqtccqc.....ctqaaaaaqtqtcaqtqaqt 554

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: qb_ba:*
2: qb_htg:*
3: qb_in:*
4: qb_om:*
5: qb_ov:*
6: qb_pat:*
7: qb_ph:*
8: qb_pl:*
9: qb_pr:*
10: qb_ro:*
11: qb_sts:*
12: qb_sy:*
13: qb_un:*
14: qb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htq_hum:*
31: em_htq_inv:*
32: em_htq_other:*
33: em_htq_mus:*
34: em_htq_pln:*
35: em_htq_rod:*
36: em_htq_mam:*
37: em_htq_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

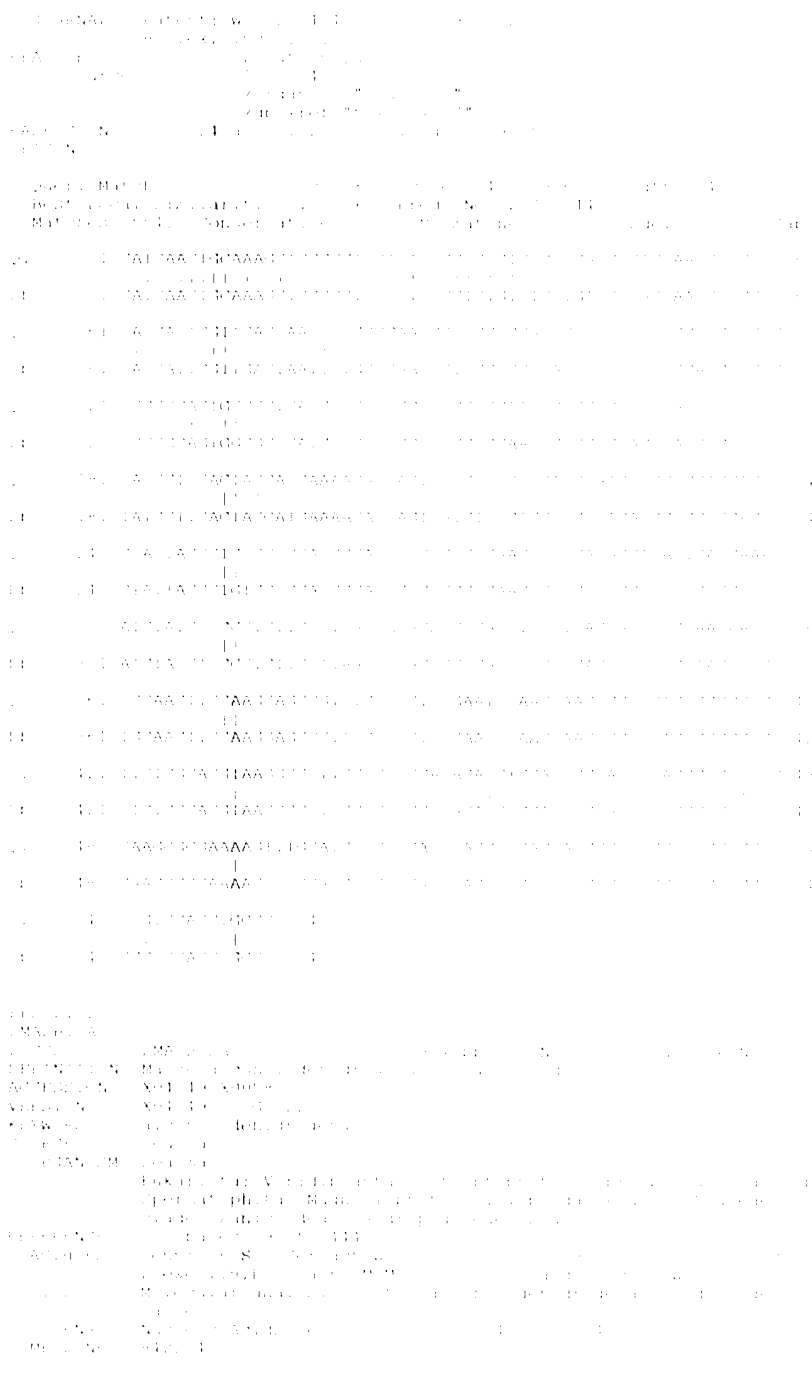
[illegible]

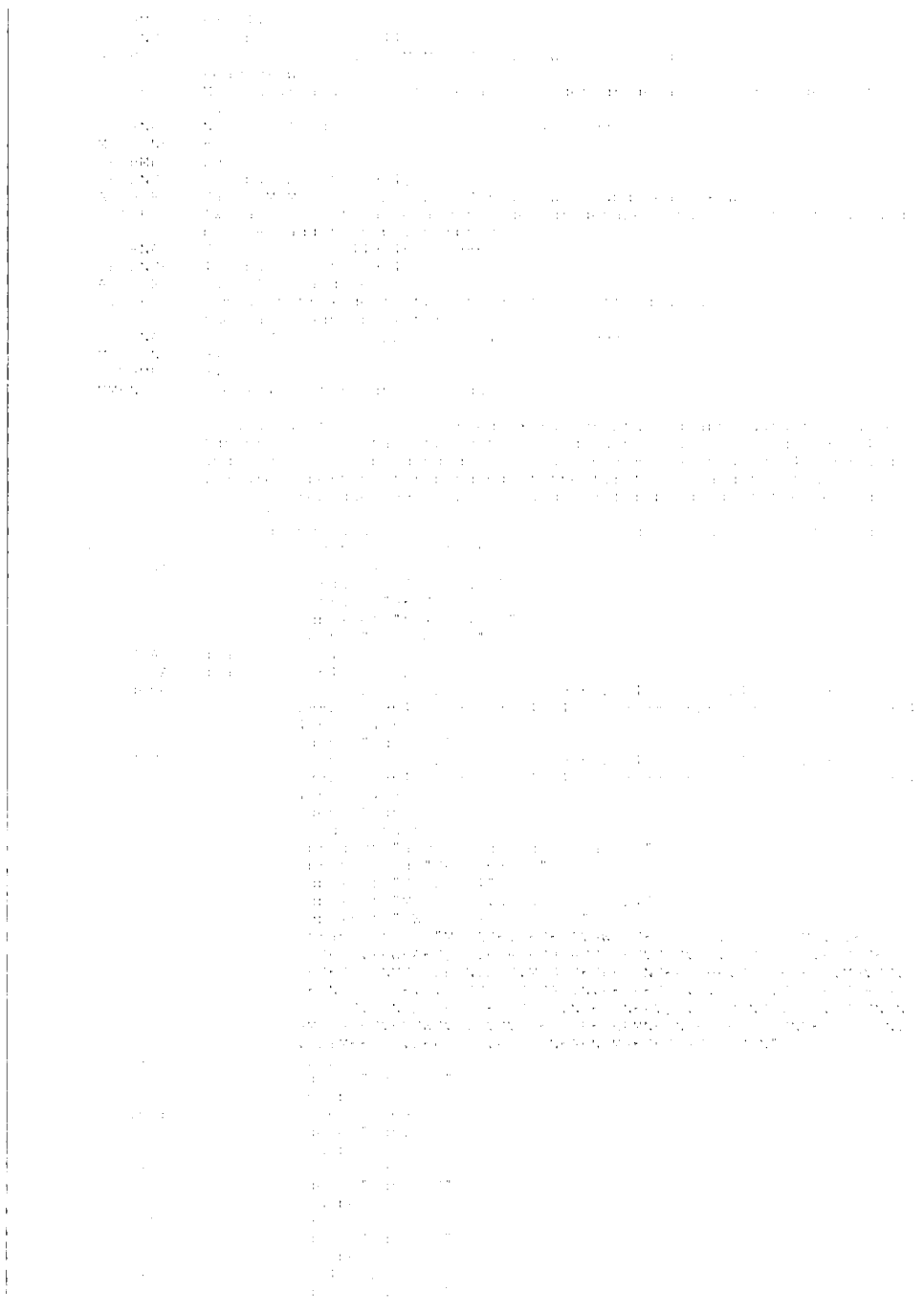
SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	554	100.0	554	1	AX[illegible]	AX[illegible]
2	554	100.0	5041	1	[illegible]	[illegible]
3	511.4	92.3	[illegible]	1	[illegible]	[illegible]
4	511.4	92.3	[illegible]	1	[illegible]	[illegible]
5	510	92.1	2722	1	[illegible]	[illegible]
6	510	92.1	2722	1	[illegible]	[illegible]
7	504	91.0	[illegible]	1	[illegible]	[illegible]
8	502	90.6	[illegible]	1	[illegible]	[illegible]
9	502	90.6	[illegible]	1	[illegible]	[illegible]
10	502	90.6	[illegible]	1	[illegible]	[illegible]
11	520.2	93.9	4194	1	[illegible]	[illegible]
12	520.2	93.9	4194	1	[illegible]	[illegible]
13	515.6	93.1	[illegible]	1	[illegible]	[illegible]
14	515.6	93.1	[illegible]	1	[illegible]	[illegible]
15	515.6	93.1	[illegible]	1	[illegible]	[illegible]
16	512	92.4	5038	1	AX[illegible]	AX[illegible]
17	512	92.4	5038	1	AX[illegible]	AX[illegible]
18	512	92.4	5038	1	AX[illegible]	AX[illegible]
19	172	31.0	454	1	AF[illegible]	AF[illegible]
20	172	31.0	459	1	AF[illegible]	AF[illegible]
21	161	29.1	849	1	AF[illegible]	AF[illegible]
22	161	29.1	850	1	AF[illegible]	AF[illegible]
23	161	29.1	860	1	AF[illegible]	AF[illegible]
24	161	29.1	860	1	AF[illegible]	AF[illegible]
25	161	29.1	861	1	AF[illegible]	AF[illegible]
26	160	28.9	456	1	AF[illegible]	AF[illegible]
27	160	28.9	457	1	AF[illegible]	AF[illegible]
28	160	28.9	857	1	AF[illegible]	AF[illegible]
29	160	28.9	857	1	AF[illegible]	AF[illegible]
30	160	28.9	857	1	AF[illegible]	AF[illegible]
31	160	28.9	857	1	AF[illegible]	AF[illegible]
32	160	28.9	857	1	AF[illegible]	AF[illegible]
33	160	28.9	857	1	AF[illegible]	AF[illegible]
34	160	28.9	857	1	AF[illegible]	AF[illegible]
35	160	28.9	857	1	AF[illegible]	AF[illegible]
36	160	28.9	857	1	AF[illegible]	AF[illegible]
37	160	28.9	857	1	AF[illegible]	AF[illegible]
38	160	28.9	857	1	AF[illegible]	AF[illegible]
39	160	28.9	857	1	AF[illegible]	AF[illegible]
40	160	28.9	857	1	AF[illegible]	AF[illegible]
41	160	28.9	857	1	AF[illegible]	AF[illegible]
42	160	28.9	858	1	AF[illegible]	AF[illegible]
43	160	28.9	858	1	AF[illegible]	AF[illegible]
44	160	28.9	859	1	AF[illegible]	AF[illegible]
45	160	28.9	859	1	AF[illegible]	AF[illegible]

ALIGNMENTS

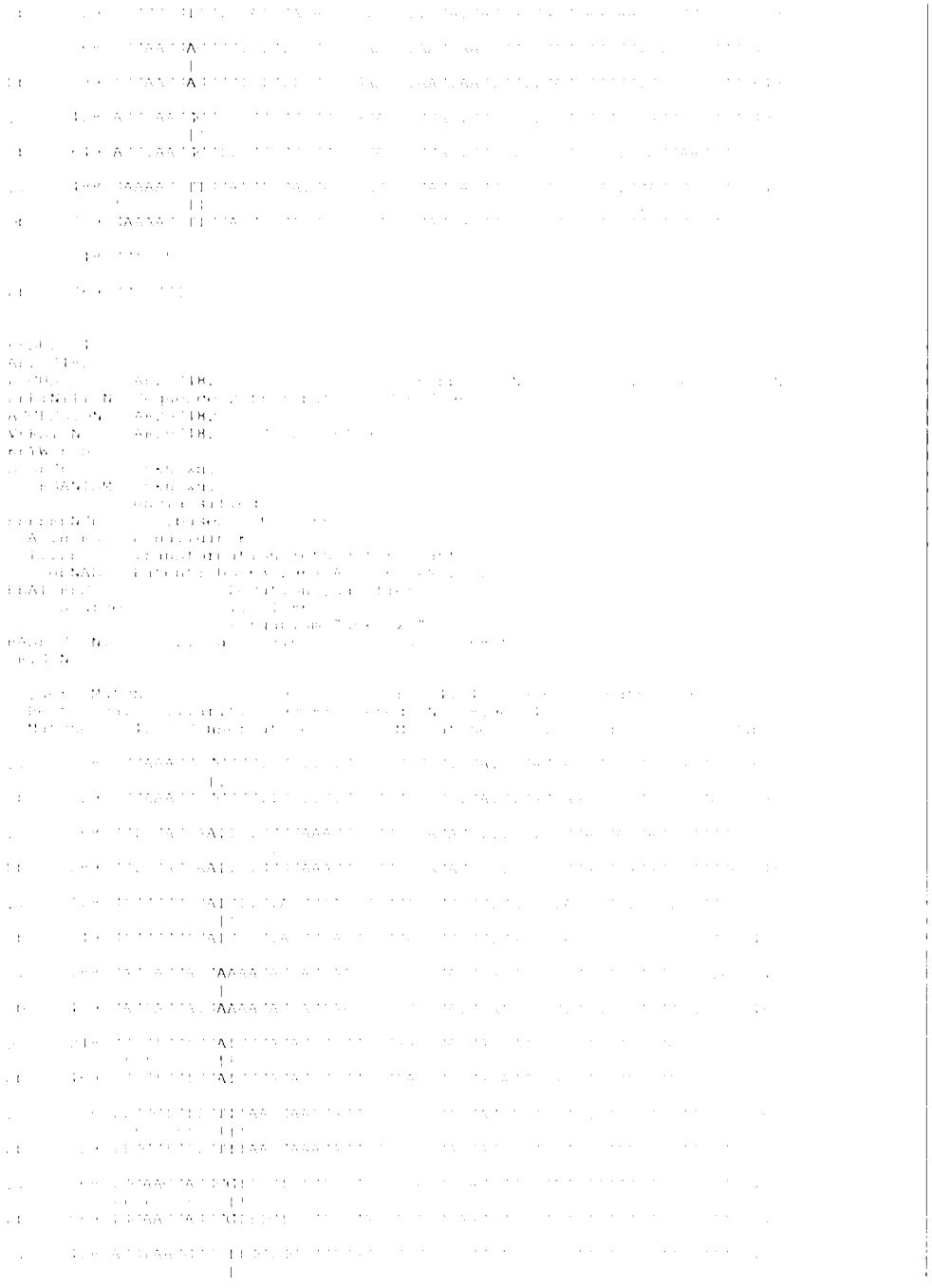
RESULT 1
AX[illegible]
LOCUS AX[illegible] [illegible] bp DNA [illegible]
DEFINITION Sequence 1 from patent W[illegible]
ACCESSION AX[illegible]
VERSION AX[illegible].1 GI:[illegible]
KEYWORDS .
SOURCE Zea mays.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Strep[illegible]
Spermatophyta; Mag[illegible]
Clade; [illegible]; Poales; [illegible]; Zea.
REFERENCE 1
AUTHORS Stomp,A.M., Dickey,L. and Siebenkas,...
TITLE Express[illegible]



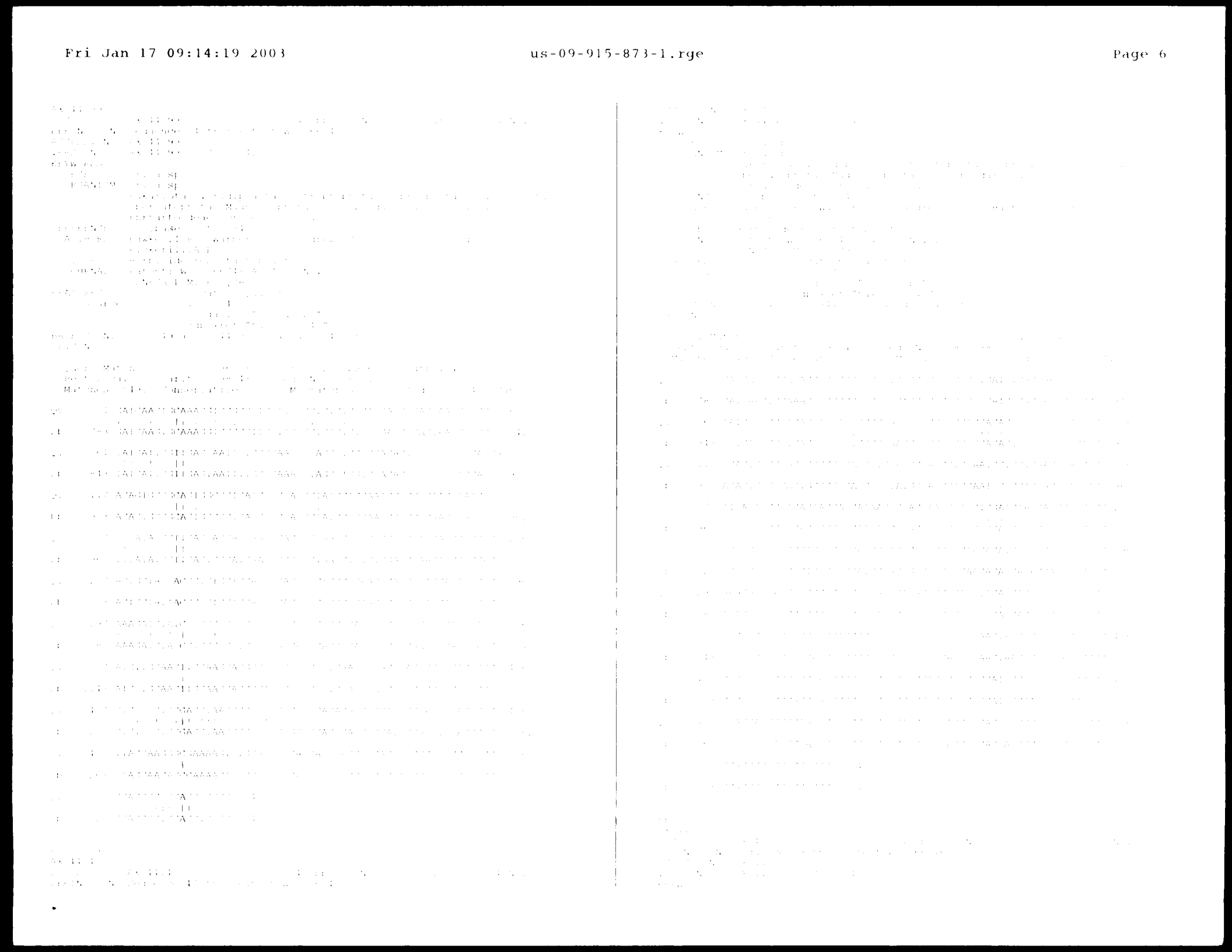


```
                     /number=3
     intron          2051..2469
                     /gene="Adh1-1S"
                     /number=3
     exon            2470..2795
                     /gene="Adh1-1S"
                     /number=4
     intron          2796..2881
                     /gene="Adh1-1S"
                     /number=4
     exon            2882..2964
                     /gene="Adh1-1S"
                     /number=5
     intron          2965..3058
                     /gene="Adh1-1S"
                     /number=5
     exon            3059..3134
                     /gene="Adh1-1S"
                     /number=6
     intron          3135..3476
                     /gene="Adh1-1S"
                     /number=6
     exon            3477..3538
                     /gene="Adh1-1S"
                     /number=7
     intron          3539..3625
                     /gene="Adh1-1S"
                     /number=7
     exon            3626..3721
                     /gene="Adh1-1S"
                     /number=8
     intron          3722..3812
                     /gene="Adh1-1S"
                     /number=8
     exon            3813..3974
                     /gene="Adh1-1S"
                     /number=9
     intron          3975..4064
                     /gene="Adh1-1S"
                     /number=9
     exon            4065..>4181
                     /gene="Adh1-1S"
                     /number=10
     polyA_signal    4223..4228
     polyA_site      4380
     polyA_site      4396
     polyA_site      4405
     polyA_site      4415
     repeat_region   4445..4448
                     /note="direct repeat"
     repeat_region   4449..4452
                     /note="direct repeat"
BASE COUNT     1520 a   1295 c   1307 g   1782 t
ORIGIN

  Query Match             100.0%;  Score 554;  DB 8;  Length 5904;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-144;
  Matches  554;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GATCAAGTGCAAAGGTCGGCCTTGTTTCTCCTCTGTCTCTTGATCTGACTAATCTTGGTT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1222 GATCAAGTGCAAAGGTCGGCCTTGTTTCTCCTCTGTCTCTTGATCTGACTAATCTTGGTT 1281

Qy         61 TATGATTCGTTGAGTAATTTTGGGGAAAGCTTCGTCCACAGTTTTTTTTTGGATGAACAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1282 TATGATTCGTTGAGTAATTTTGGGGAAAGCTTCGTCCACAGTTTTTTTTTGGATGAACAG 1341

Qy        121 TGCCGCAGTGGCGCTGATCTTGTATGCTATCCTGCAATCGTGGTGAACTTATGTCTTTTA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1342 TGCCGCAGTGGCGCTGATCTTGTATGCTATCCTGCAATCGTGGTGAACTTATGTCTTTTA 1401

Qy        181 TATCCTTCACTACCATGAAAAGACTAGTAATTTTTCTGGA[illegible] 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

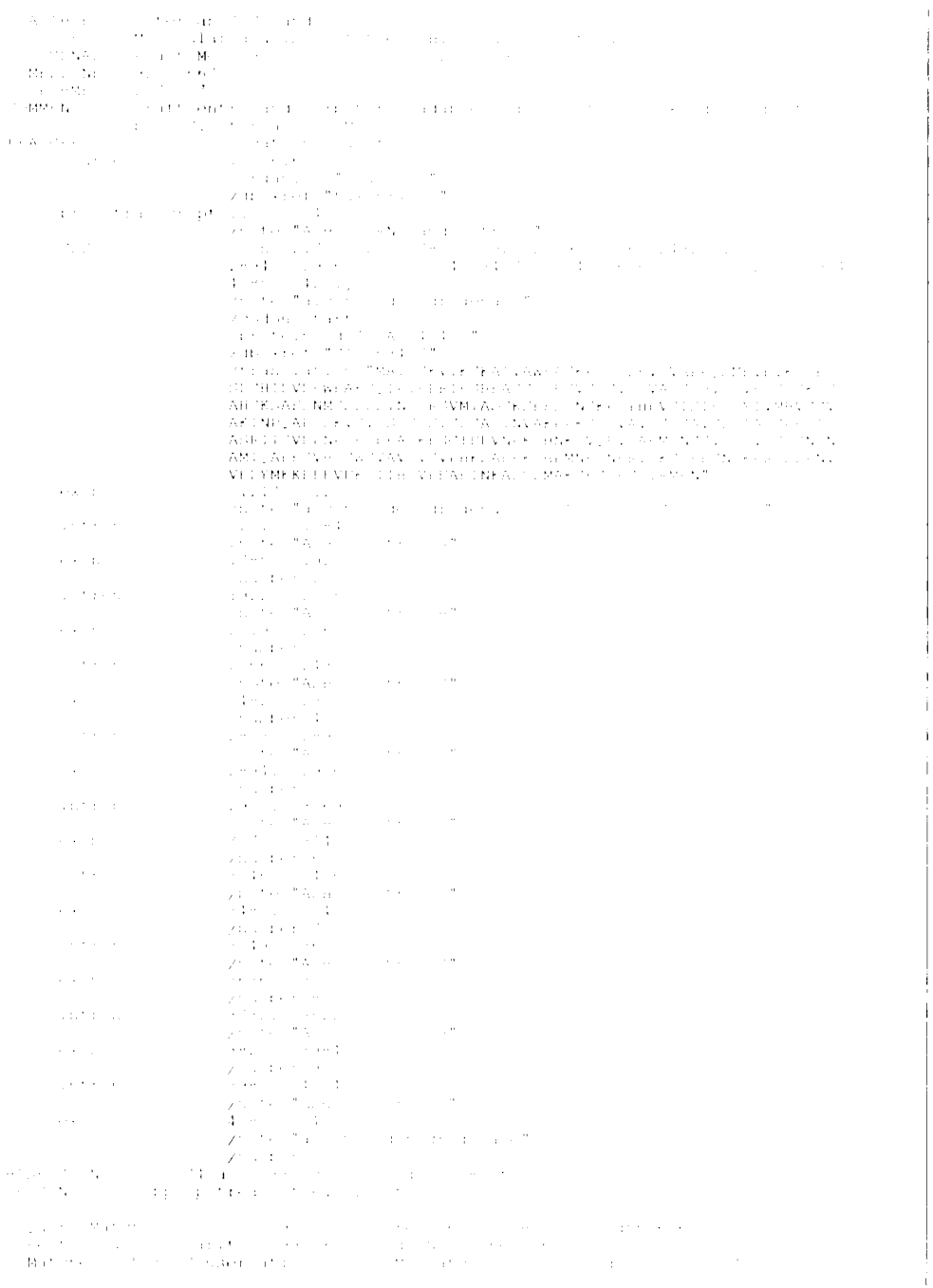
```
Db       1402 [illegible]

Qy        241 [illegible]
Db       1462 [illegible]

Qy        301 [illegible]
Db       1522 [illegible]

Qy        361 [illegible]
Db       1582 [illegible]

Qy        421 [illegible]
Db       1642 [illegible]

Qy        481 [illegible]
Db       1702 [illegible]

Qy        541 [illegible]
Db       1762 [illegible]


RESULT [illegible]
[illegible]
LOCUS       [illegible]      DNA     [illegible]
DEFINITION  Sequence 2 from patent [illegible]
ACCESSION   [illegible]
VERSION     [illegible]
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to [illegible])
  AUTHORS   D'Halluin,K.
  TITLE     Transformation [illegible]
  JOURNAL   Patent: US [illegible]
FEATURES             Location/Qualifiers
     source          1..[illegible]
                     /organism="unknown"
BASE COUNT     [illegible]
ORIGIN

  Query Match             [illegible]
  Best Local Similarity   [illegible];  Pred. No. [illegible];
  Matches  542;  Conservative    [illegible];  Mismatches    [illegible];  Indels    [illegible];  Gaps    [illegible]

Qy          8 [illegible]
Db        229 [illegible]

Qy         68 [illegible]
Db        289 [illegible]

Qy        128 [illegible]
Db        349 [illegible]

Qy        188 [illegible]
Db        409 [illegible]

Qy        248 [illegible]
Db        469 [illegible]

Qy        308 [illegible]
```



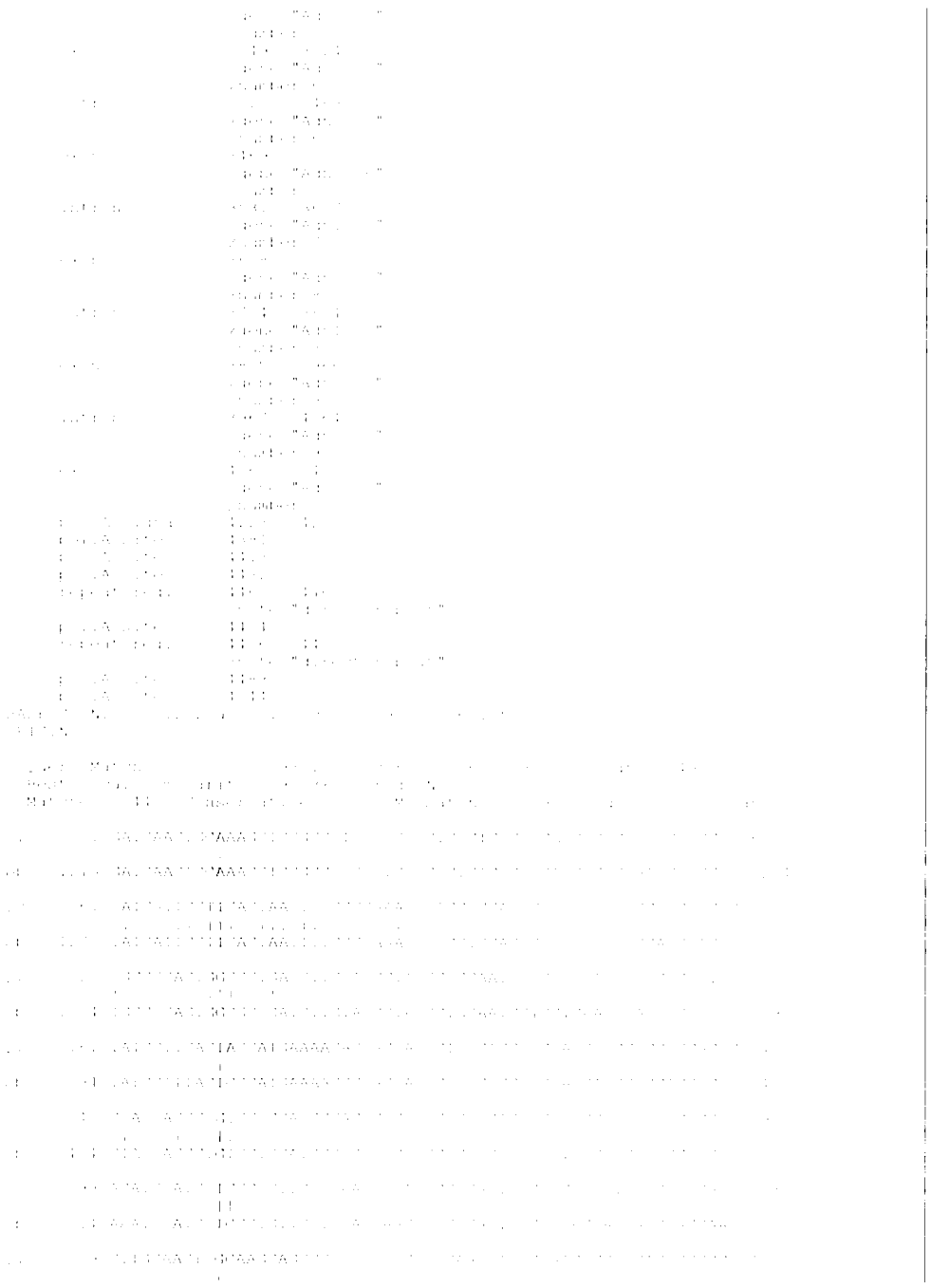
```
Qy     477 TTAGCAAGGGCGAAAAGTTTGCATCTTGATGATTTAGCTTGACTATGCGATTGCTTTCCT 536
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1436 TTAGCAAGGGCGAAAAGTTTGCATCTTGATGATTTAGCTTGACTATGCGATTGCTTTCCT 1377

Qy     537 GGACCCGTGCAGCTGCGG 554
           ||||||||||||||||||
Db    1496 GGACCCGTGCAGCTGCGG 1513


RESULT 6
AX039931
LOCUS       AX039931               2722 bp    DNA     linear   PAT 18-N-V-2000
DEFINITION  Sequence 9 from Patent WO0063401.
ACCESSION   AX039931
VERSION     AX039931.1  GI:11229960
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences
REFERENCE   1  (bases 1 to 2722)
  AUTHORS   Habben,J.E., Zinselmeier,C. and Tomes,D.
  TITLE     Regulated expression of genes in plant seeds
  JOURNAL   Patent: WO 0063401-A 9 26-OCT-2000;
            PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1. .2722
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Promoter from Zea mays, U.S. patent application
                     09/377,648; gene from Arabidopsis thaliana, Mol. & Gen.
                     and General Genetics 216:388-394 (1989); terminator from
                     Solanum tuberosum, Plant Cell 1(1):115-122 (1989)."
BASE COUNT      732 a    610 c    619 g    761 t
ORIGIN

  Query Match             97.5%;  Score 540;  DB 6;  Length 2722;
  Best Local Similarity   99.3%;  Pred. No. 2.2e-140;
  Matches  554;  Conservative    0;  Mismatches    4;  Indels    3;  Gaps    3;

Qy       1 GATCAAGTGCAAAGGTCCGGCTTGTTCTCCTCTGTCTCTTGATCTGACTAAATCTTGATT 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     956 GATCAAGTGCAAAGGTCCGGCTTGTTCTCCTCTGTCTCTTGATCTGACTAAATCTTGATT 897

Qy      61 TATGATTCGTTGACTAATTTTGGGGAAAGCTTCGTCACAGTTTTTTTTTTGAGAAAAAT 12
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1016 TATGATTCGTTGACTAATTTTGGGGAAAGCTTCGTCACAGTTTTTTTTTTGAGAAAAAT 957

Qy     121 TGCCGCAGTGGCGCTGATCTTGTATGCTATCTTGCAAACGTGTCTGAATTTATTCTTTTA 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1076 TGCCGCAGTGGCGCTGATCTTGTATGCTATCTTGCAAACGTGTCTGAATTTATTCTTTTA 1017

Qy     181 TATCCTTCACTACCATGAAAAGACTAGTAATCTTTCTCGCATGTAACTATCGTCACTATCG 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1136 TATCCTTCACTACCATGAAAAGACTAGTAATCTTTCTCGCATGTAACTATCGTCACTATCG 1077

Qy     241 CTATTACCGTGTGCTCGATCCGACAGTCTGCTGAACACATCATACTATATTGACAAAAG 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1196 CTATTACCGTGTGCTCGATCCGACAGTCTGCTGAACACATCATACTATATTGACAAAAG 1137

Qy     300 ---GATCTATCTTCCCTGTTCTTTAATGAAAGACGTCATTTCATCATATAATAATAAGA 356
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1256 ATCGATCTATCTTCCCTGTTCTTTAATGAAAGACGTCATTTCATCATATAATAATAAGA 1197

Qy     357 ATGTTGCAACTTGCAAGGAGGCGTTTCTTTCTTTGAATTTAACTAACTCGTTGAGTGGGT 416
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1316 ATGTTGCAACTTGCAAGGAGGCGTTTCTTTCTTTGAATTTAACTAACTCGTTGAGTGGGT 1257

Qy     417 CTGTTTCTTGGGACGTAAGGCCTTTGCTGCTCCACACATGTCCATTCTAACTTTACTTCT 476
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1376 CTGTTTCTTGGGACGTAAGGCCTTTGCTGCTCCACACATGTCCATTCTAACTTTACTTCT 1317

Qy     477 TTAGCAAGGGCGAAAAGTTTGCATCTTGATGATTTAGCTTGACTATGCGATTGCTTTCCT 536
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1436 TTAGCAAGGGCGAAAAGTTTGCATCTTGATGATTTAGCTTGACTATGCGATTGCTTTCCT 1377

Qy     537 GGACCCGTGCAGCTGCGG 554
           ||||||||||||||||||
Db    1496 GGACCCGTGCAGCTGCGG 1513


RESULT 7
AX039931
LOCUS       AX039931               2722 bp    DNA     linear   PAT 18-N-V-2000
DEFINITION  Sequence 9 from Patent WO0063401.
ACCESSION   AX039931
VERSION     AX039931.1  GI:11229960
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences
REFERENCE   1  (bases 1 to 2722)
  AUTHORS   Habben,J.E., Zinselmeier,C. and Tomes,D.
  TITLE     Regulated expression of genes in plant seeds
  JOURNAL   Patent: WO 0063401-A 9 26-OCT-2000;
            PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1. .2722
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Promoter from Zea mays, U.S. patent application
                     09/377,648; gene from Arabidopsis thaliana, Mol. & Gen.
                     and General Genetics 216:388-394 (1989); terminator from
                     Solanum tuberosum, Plant Cell 1(1):115-122 (1989)."
BASE COUNT      732 a    610 c    619 g    761 t
ORIGIN

  Query Match             97.5%;  Score 540;  DB 6;  Length 2722;
  Best Local Similarity   99.3%;  Pred. No. 2.2e-140;
  Matches  554;  Conservative    0;  Mismatches    4;  Indels    3;  Gaps    3;

Qy       1 GATCAAGTGCAAAGGTCCGGCTTGTTCTCCTCTGTCTCTTGATCTGACTAAATCTTGATT 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1277 GATCAAGTGCAAAGGTCCGGCTTGTTCTCCTCTGTCTCTTGATCTGACTAAATCTTGATT 1218

Qy      61 TATGATTCGTTGACTAATTTTGGGGAAAGCTTCGTCACAGTTTTTTTTTTGAGAAAAAT 12
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1337 TATGATTCGTTGACTAATTTTGGGGAAAGCTTCGTCACAGTTTTTTTTTTGAGAAAAAT 1278

Qy     121 TGCCGCAGTGGCGCTGATCTTGTATGCTATCTTGCAAACGTGTCTGAATTTATTCTTTTA 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1397 TGCCGCAGTGGCGCTGATCTTGTATGCTATCTTGCAAACGTGTCTGAATTTATTCTTTTA 1338

Qy     181 TATCCTTCACTACCATGAAAAGACTAGTAATCTTTCTCGCATGTAACTATCGTCACTATCG 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1457 TATCCTTCACTACCATGAAAAGACTAGTAATCTTTCTCGCATGTAACTATCGTCACTATCG 1398

Qy     241 CTATTACCGTGTGCTCGATCCGACAGTCTGCTGAACACATCATACTATATTGACAAAAG 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1517 CTATTACCGTGTGCTCGATCCGACAGTCTGCTGAACACATCATACTATATTGACAAAAG 1458

Qy     300 ---GATCTATCTTCCCTGTTCTTTAATGAAAGACGTCATTTCATCATATAATAATAAGA 356
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1577 ATCGATCTATCTTCCCTGTTCTTTAATGAAAGACGTCATTTCATCATATAATAATAAGA 1518

Qy     357 ATGTTGCAACTTGCAAGGAGGCGTTTCTTTCTTTGAATTTAACTAACTCGTTGAGTGGGT 416
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1637 ATGTTGCAACTTGCAAGGAGGCGTTTCTTTCTTTGAATTTAACTAACTCGTTGAGTGGGT 1578

Qy     417 CTGTTTCTTGGGACGTAAGGCCTTTGCTGCTCCACACATGTCCATTCTAACTTTACTTCT 476
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1697 CTGTTTCTTGGGACGTAAGGCCTTTGCTGCTCCACACATGTCCATTCTAACTTTACTTCT 1638

Qy     477 TTAGCAAGGGCGAAAAGTTTGCATCTTGATGATTTAGCTTGACTATGCGATTGCTTTCCT 536
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1757 TTAGCAAGGGCGAAAAGTTTGCATCTTGATGATTTAGCTTGACTATGCGATTGCTTTCCT 1698

Qy     537 GGACCCGTGCAGCTGCGG 554
           ||||||||||||||||||
Db    1817 GGACCCGTGCAGCTGCGG 1834


RESULT 8
```



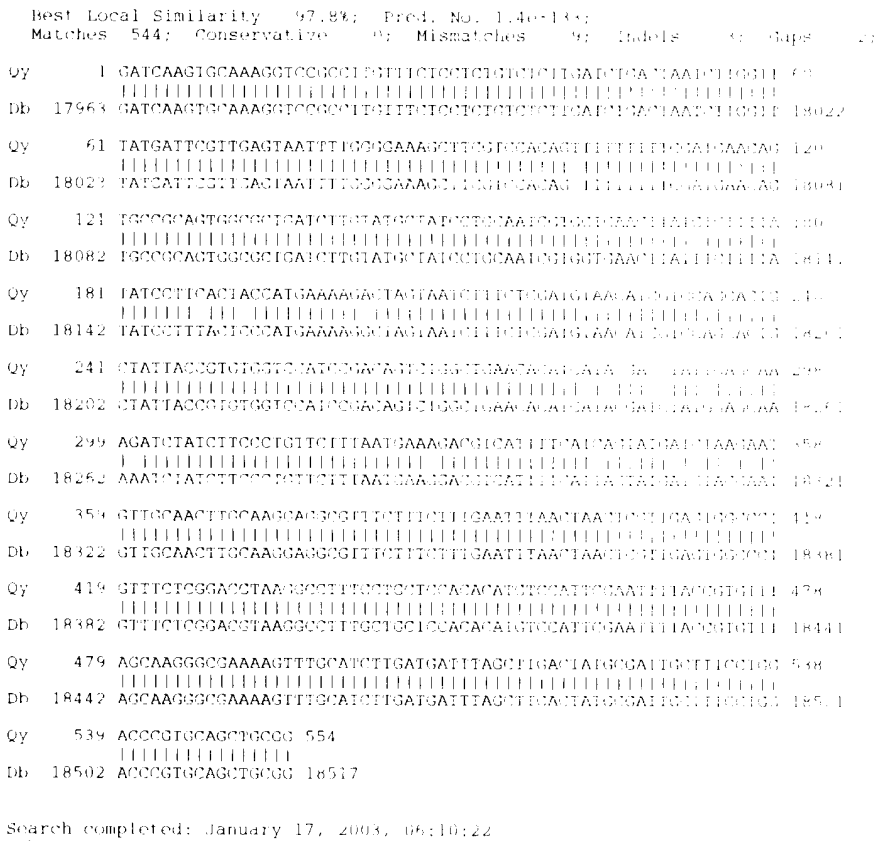
```
SOURCE      Oryza sp.
  ORGANISM  Oryza sp.
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 1343)
  AUTHORS   Hawkes,T.R., Warner,S.A., Andrews,C.J., Bachoo,S. and
            Pickerill,A.P.
  TITLE     Herbicide resistant plants
  JOURNAL   Patent: WO 0066746 A 9 09 NOV 2000;
            ZENECA LIMITED (GB)
FEATURES             Location/Qualifiers
     source          1. .1343
                     /organism="Oryza sp."
                     /db_xref="taxon:52841"
BASE COUNT      349 a    344 c    245 g    465 t
ORIGIN

  Query Match             96.0%;  Score 532;  DB 6;  Length 1343;
  Best Local Similarity   96.4%;  Pred. No. 7.8e-139;
  Matches  549;  Conservative    0;  Mismatches    5;  Indels    4;  Gaps    1;

Qy       1 GATCAAGTGCAAAGGTCCGCCTIGTTTCTCCTCTGTCTCTTGATCTGATTAATTTTGGTT 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     783 GATCAAGTGCAAAGGTCCGCCTIGTTTCTCCTCTGTCTCTTGATCTGATTAATTTTGGTT 842

Qy      61 TATGATTCGTTGAGTAATTTTGGGGAA----AGCTTCGTCCACAGTTTTTTTTTCGAIGA 116
           |||||||||||||||||||||||||||    ||||||||||||||| ||||||||||||
Db     843 TATGATTCGTTGAGTAATTTTGGGGAAAGCTAGCTTCGTCCACAGTTTTTTTTTCGAIGA 902

Qy     117 ACACTGCCGCAGTGGCGCTGATCTTGTATGCTATCCTGCAATCCTGGTGGAACTTATGTCT 176
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     903 ACAGTGCCGCAGTGGCGCTGATCTTGTATGCTATCCTGCAATCCTGGTGGAACTTATTTCT 962

Qy     177 TTTATATCCTTCACTACCATGAAAAGACTAGTAATCTTTCTCGATGTAACTAATCATCGTCA 236
           |||||||||||||||| |||||||||| ||||||||||||||||||||||||||||||||
Db     963 TTTATATCCTTCACTCCCATGAAAAGGCTAGTAATCTTTCTCGATGTAACTAATCATCGTCA 1022

Qy     237 ACTGCTATTACCGTGTGGTCCATCCGACAGTCTGGCTGAACACATCATACTATTGAGC 296
           ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db    1023 ACTGCTATTACCGTGTGGTCCATCCGACAGTCTGGCTGAACACATCATACTATTGAGC 1082

Qy     297 AAAGATCTATCTTCCCTGTTCTTTAATGAAAGACGTCATTTTCATCAGTATGATCTAATGA 356
           |||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db    1083 AAAGATCTATCCTCCCTGTTCTTTAATGAAAGACGTCATTTTCATCAGTATGATCTAATGA 1142

Qy     357 ATGTTGCAACTTGCAAGGAGGCGTTTCTTTCTTTGAATTTAACTAACTCGTTGAGTGGCC 416
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1143 ATGTTGCAACTTGCAAGGAGGCGTTTCTTTCTTTGAATTTAACTAACTCGTTGAGTGGCC 1202

Qy     417 CTGTTTCTCGGACGTAAGGCCTTTGCTGCTCCACACATGTCTATTCAAATTTAACGTTCT 476
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1203 CTGTTTCTCGGACGTAAGGCCTTTGCTGCTCCACACATGTCTATTCAAATTTAACGTTCT 1262

Qy     477 TTAGCAAGGGCGAAAAGTTTGCATCTTGATGATTTAGCTTGACTATGGATTTTTTCTTCT 536
           |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1263 TTAGCAAGAGCGAAAAGTTTGCATCTTGATGATTTAGCTTGACTATGGATTTTTTCTTCT 1322

Qy     537 GGACCCGTGCAGCTGCGG 554
           ||||||||||||||||||
Db    1323 GGACCCGTGCAGCTGCGG 1340


RESULT 11
ZMADH1ALL
LOCUS       ZMADH1ALL               6158 bp    DNA     linear   PLN 23-JAN-1996
DEFINITION  Z.mays DNA for Adh1-Cm allele.
ACCESSION   X17556
VERSION     X17556.1  GI:22118
KEYWORDS    ADH1 gene.
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 6158)
  AUTHORS   Dennis,E.S. and Llewellyn,D.
  TITLE     Molecular analysis of the ADH1 Cm allele of maize
  JOURNAL   Plant Mol. Biol. 14 (2), 253-254 (1990)
  MEDLINE   92063667
   PUBMED   2277507
FEATURES             Location/Qualifiers
     source          1. .6158
                     /organism="Zea mays"
                     /db_xref="taxon:4577"
BASE COUNT     1971 a   1355 c   1387 g   1445 t
ORIGIN

  Query Match             96.0%;  Score 532;  DB 6;  Length 6158;
  Best Local Similarity   96.4%;  Pred. No. 3.4e-138;
  Matches  549;  Conservative    0;  Mismatches    5;  Indels    4;  Gaps    1;

Qy       1 GATCAAGTGCAAAGGTCCGCCTIGTTTCTCCTCTGTCTCTTGATCTGATTAATTTTGGTT 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1228 GATCAAGTGCAAAGGTCCGCCTIGTTTCTCCTCTGTCTCTTGATCTGATTAATTTTGGTT 1287

Qy      61 TATGATTCGTTGAGTAATTTTGGGGAA----AGCTTCGTCCACAGTTTTTTTTTCGAIGA 116
           ||||||||||||||||||||||||||||    ||||||||||||||| ||||||||||||
Db    1288 TATGATTCGTTGAGTAATTTTGGGGAAAGCTAGCTTCGTCCACAGTTTTTTTTTCGAIGA 1347

Qy     120 GTGCCGCAGTGGCGCTGATCTTGTATGCTATCCTGCAATCCTGGTGGAACTTATGTCTTTT 179
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1348 GTGCCGCAGTGGCGCTGATCTTGTATGCTATCCTGCAATCCTGGTGGAACTTATTTCTTTT 1407

Qy     180 ATATCCTTCACTACCATGAAAAGACTAGTAATCTTTCTCGATGTAACTAATCATCGTCAAC 239
           |||||||||| |||||||||| |||||||||||||||||||||||||||||||||||||||
Db    1408 ATATCCTTCACTCCCATGAAAAGGCTAGTAATCTTTCTCGATGTAACTAATCATCGTCAAC 1467

Qy     240 TGCTATTACCGTGTGGTCCATCCGACAGTCTGGCTGAACACATCATACTATTGAGCAAAG 299
           |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db    1468 TGCTATTACCGTGTGGTCCATCCGACAGTCTGGCTGAACACATCATACTATTGAGCAAAG 1527

Qy     300 ATCTATCTTCCCTGTTCTTTAATGAAAGACGTCATTTTCATCAGTATGATCTAATGAATG 359
           |||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1528 ATCTATCCTCCCTGTTCTTTAATGAAAGACGTCATTTTCATCAGTATGATCTAATGAATG 1587

Qy     360 TTGCAACTTGCAAGGAGGCGTTTCTTTCTTTGAATTTAACTAACTCGTTGAGTGGCCCTG 419
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1588 TTGCAACTTGCAAGGAGGCGTTTCTTTCTTTGAATTTAACTAACTCGTTGAGTGGCCCTG 1647

Qy     420 TTTCTCGGACGTAAGGCCTTTGCTGCTCCACACATGTCTATTCAAATTTAACGTTCTTTA 479
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1648 TTTCTCGGACGTAAGGCCTTTGCTGCTCCACACATGTCTATTCAAATTTAACGTTCTTTA 1707

Qy     480 GCAAGGGCGAAAAGTTTGCATCTTGATGATTTAGCTTGACTATGGATTTTTTCTTCTGGA 539
           |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1708 GCAAGAGCGAAAAGTTTGCATCTTGATGATTTAGCTTGACTATGGATTTTTTCTTCTGGA 1767

Qy     540 CCCGTGCAGCTGCGG 554
           |||||||||||||||
Db    1768 CCCGTGCAGCTGCGG 1782


RESULT 12
M13440:1CM
LOCUS       M13440:1CM              [illegible] bp    DNA     linear   PLN [illegible]
DEFINITION  Z.mays alcohol dehydrogenase (ADH [illegible]), [illegible]
ACCESSION   M32084
VERSION     M32084.1  GI:[illegible]
KEYWORDS    alcohol dehydrogenase.
SOURCE      Z.mays DNA.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to [illegible])
```



```
     gene            <1. .2988
                     /gene="adh1"
                     /allele "adh1-F"
     mRNA            join(<1. .34,570. .706,804. .850,1266. .1591,1678. .1760,
                     1853. .1928,2274. .2435,2423. .2518,2613. .2771,2872. .2988)
                     /gene="adh1"
                     /product="alcohol dehydrogenase 1"
     CDS             join(1. .34,570. .706,804. .850,1266. .1591,1678. .1760,
                     1853. .1928,2274. .2435,2423. .2518,2613. .2771,2872. .2988)
                     /gene="adh1"
                     /note="ADH1-F"
                     /codon_start=1
                     /product="alcohol dehydrogenase 1"
                     /protein_id="AAC34295.1"
                     /db_xref="GI:3420021"
                     /translation="MATAGKVIKCKAAVAWEAGKPLSIEEVEVAPPKAMEVRVKILFT
                     SLCHTDVYFWEAKGQTPVFPRIFGHEAGGIVESVGEGVTLVAPGDHVLPVFTGECKE
                     AHCKSAESNMCDLLRINTERGVMIADGKSRFSINGKPIYHFVGTSTFSEYTVVHIGCV
                     AKINPQAPLDKVCVLSCGISTGLGAAINVAKPKKGSTVAVFGLGAVGLAAAEGARIAG
                     ASRIIGVDLNPSRFEEAKEFGCTEFVNPKDHDKPVQEVIAEMTNGGVDFSFEVIGRLD
                     AMIQAFECVHDGWGVAVLVGVPHKDAEFKTHPMNFLNERCLKGTFFGNYKPRTDLPNV
                     VELYMKKELEVEKFITHSVPFAEINKAFDLMAKGDGTRCILTMEN"
BASE COUNT      680 a    638 c    725 g    945 t
ORIGIN

  Query Match             93.1%;  Score 515.6;  DB 8;  Length 2988;
  Best Local Similarity   97.8%;  Pred. No. 1.5e-133;
  Matches  544;  Conservative    0;  Mismatches    9;  Indels    3;  Gaps    1;

Qy       1 GATCAAGTGCAAAGGTCCGGCTTGTTTCTCCTCTGTCTCTCTTGATCTGACTAAATCTTGGTT 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      21 GATCAAGTGCAAAGGTCCGGCTTGTTTCTCCTCTGTCTCTCTTGATCTGACTAAATCTTGGTT 80

Qy      61 TATGATTCGTTGAGTAATTTTGGGCAAAAGCTTCGTCCACAGTTTTTTTTTCCAATCAATATA 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      81 TATGATTCGTTGACTAATTTTGGGCAAAAGCTTCGTCCACAGTTTTTTTTTCAATCAATATA 140

Qy     121 TGCCGCAGTGGGGCTGATCTTGTATGCTATCGTCGCAATCGTGGTGAAGTTIATTCTTTTIA 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     140 TGCCGCAGTGGGGCTGATCTTGTATGCTATCGTCGCAATCGTGGTGAAGTTIATTCTTTTIA 199

Qy     181 TATCTTTCACTACCATGAAAAGACTAGTAATCTTTCTTCTATATGTAACAATCATTATAGTTA 240
           ||||||| ||| |||||||||| ||||||||||||||||||||||||||||||||||||||||
Db     200 TATCTTTACTCCCATGAAAAGGTTAGTAATCTTTCTTCTATATGTAACAATCATTATAGTTA 259

Qy     241 CTATTACCGTCGTGGTCCATCCGACAGTCTGGCTCAACACATCACTACGA TATTGATAA 299
           |||||||||||||||||||||||||||||||||||||||||||||||||| || ||||||
Db     260 CTATTACCGTCGTGGTCCATCCGACAGTCTGGCTCAACACATCACTACGAT TATCGATAA 319

Qy     299 AGATCTATCTTCCCTGTTCTTTAATCAAAGACTTCGATTTTCATTAGTATCATCTAAAAAT 358
           | |||||||||||||||||||||||||| |||||||||||| |||||||||| |||||||
Db     320 AAATCTATCTTCCCTGTTCTTTAATCAAAGACTTCGATTTTCATTAGTATCATCTAAAAAT 379

Qy     359 GTTGCAACTTGCAACGACGCGTTTCTTTCTTTGAATTTAACTAACTCGTTGACTGGGCCT 418
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     380 GTTGCAACTTGCAAGGAGGCGTTTCTTTCTTTGAATTTAACTAACTCGTTGACTGGGCCT 439

Qy     419 GTTTCTCGGACGTAAGGCCTTTGCTGCTCCACACATGTCCACTCGAATTTTATGTCTCTT 478
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     440 GTTTCTCGGACGTAAGGCCTTTGCTGCTCCACACATGTCCACTCGAATTTTATGTCTCTT 499

Qy     479 AGCAAGGGCGAAAAGTTTGCATCTTGATGATTTAGCTTGACTATGCAATTGCTTTGCTGC 538
           |||||||||||||||||||||||||||||||||||||||||||| ||||||.||||||||||
Db     500 AGCAAGGGCGAAAAGTTTGCATCTTGATGATTTAGCTTGACTATGCGATTGCTTTGCTGC 559

Qy     539 ACCCGTGCAGCTGCGG 554
           ||||||||||||||||
Db     560 ACCCGTGCAGCTGCGG 575


RESULT 14
ZMADH1FA
LOCUS       ZMADH1FA                5249 bp    DNA     linear   PLN 07-Aug-1995
DEFINITION  Maize alcohol dehydrogenase 1 gene (Adh1-F)
ACCESSION   X04049
VERSION     X04049.1  GI:22...
KEYWORDS    alcohol dehydrogenase.
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 5249)
  AUTHORS   Sachs,M.M., Freeling,M.S., Okimoto,R. and Borner,T.K.
  TITLE     Two alleles of maize alcohol dehydrogenase 1 have ... and
            poly(A) addition ...
  JOURNAL   Submitted (13-AUG-1987)
COMMENT     See X04048 for Adh1-S allele
FEATURES             Location/Qualifiers
     source          1..5249
                     /organism="Zea mays"
                     /strain "B68"
                     /db_xref="taxon:4577"
                     /clone_lib "Lambda EMBL4"
                     /clone "Adh1-F"
     CAAT_signal     997..1000
     TATA_signal     1065..1072
     mRNA            join(1..34,570..706,804..850,1266..1591,1678..1760,
                     1853..1928,2274..2435,2423..2518,2613..2771,2872..2988)
                     /gene="adh1"
                     /allele "adh1-F"
     CDS             join(1..34,570..706,804..850,1266..1591,1678..1760,
                     1853..1928,2274..2435,2423..2518,2613..2771,2872..2988)
                     /gene="adh1"
                     /codon_start=1
                     /product="alcohol dehydrogenase 1"
                     /protein_id="AAA..."
                     /db_xref="GI:...."
                     /db_xref="Mutation/Variation"
                     /db_xref="SWISS-PROT:..."
                     /translation="MATAGKVIKCKAAVAWEAGKPLSIEEVEVAPPKAMEVRVKILFT
                     SLCHTDVYFWEAKGQTPVFPRIFGHEAGGIVESVGEGVTLVAPGDHVLPVFTGECKE
                     AHCKSAESNMCDLLRINTERGVMIADGKSRFSINGKPIYHFVGTSTFSEYTVVHIGCV
                     AKINPQAPLDKVCVLSCGISTGLGAAINVAKPKKGSTVAVFGLGAVGLAAAEGARIAG
                     ASRIIGVDLNPSRFEEAKEFGCTEFVNPKDHDKPVQEVIAEMTNGGVDFSFEVIGRLD
                     AMIQAFECVHDGWGVAVLVGVPHKDAEFKTHPMNFLNERCLKGTFFGNYKPRTDLPNV
                     VELYMKKELEVEKFITHSVPFAEINKAFDLMAKGDGTRCILTMEN"
     exon            1..34
                     /gene="Adh1-F"
                     /number=1
     intron          1..
                     /gene="Adh1-F"
                     /number=1
     exon            570..706
                     /gene="Adh1-F"
                     /number=2
     intron          707..
                     /gene="Adh1-F"
                     /number=2
     exon            804..850
                     /gene="Adh1-F"
                     /number=3
     intron          851..1265
                     /gene="Adh1-F"
                     /number=3
     exon            1266..1591
                     /gene="Adh1-F"
                     /number=4
     intron          1592..1677
                     /gene="Adh1-F"
                     /number=4
     exon            1678..1760
                     /gene="Adh1-F"
                     /number=5
     intron          1761..1852
                     /gene="Adh1-F"
                     /number=5
     exon            1853..1928
                     /gene="Adh1-F"
                     /number=6
     intron          1929..2273
                     /gene="Adh1-F"
                     /number=6
     exon            2274..2435
                     /gene="Adh1-F"
                     /number=7
     intron          2436..2422
                     /gene="Adh1-F"
                     /number=7
     exon            2423..2518
                     /gene="Adh1-F"
                     /number=8
     intron          2519..2612
                     /gene="Adh1-F"
                     /number=8
     exon            2613..2771
                     /gene="Adh1-F"
                     /number=9
     intron          2772..2871
                     /gene="Adh1-F"
                     /number=9
     exon            2872..2988
                     /gene="Adh1-F"
                     /number=10
     polyA_signal    3000..3005
     polyA_signal    3020..3025
     polyA_signal    3040..3045
     polyA_signal    3060..3065
     polyA_signal    3080..3085
     polyA_signal    3100..3105
     polyA_signal    3120..3125
     polyA_signal    3140..3145
     polyA_signal    3160..3165
     polyA_signal    3180..3185
     polyA_signal    3200..3205
     polyA_signal    3220..3225
     polyA_signal    3240..3245
     polyA_signal    3260..3265
     polyA_signal    3280..3285
     polyA_signal    3300..3305
     polyA_signal    3320..3325
     polyA_signal    3340..3345
     polyA_signal    3360..3365
     polyA_signal    3380..3385
     polyA_signal    3400..3405
     polyA_signal    3420..3425
     polyA_signal    3440..3445
     polyA_signal    3460..3465
     polyA_signal    3480..3485
     polyA_signal    3500..3505
     polyA_signal    3520..3525
     polyA_signal    3540..3545
     polyA_signal    3560..3565
     polyA_signal    3580..3585
     polyA_signal    3600..3605
     polyA_signal    3620..3625
     polyA_signal    3640..3645
     polyA_signal    3660..3665
     polyA_signal    3680..3685
     polyA_signal    3700..3705
     polyA_signal    3720..3725
     polyA_signal    3740..3745
     polyA_signal    3760..3765
     polyA_signal    3780..3785
     polyA_signal    3800..3805
     polyA_signal    3820..3825
     polyA_signal    3840..3845
     polyA_signal    3860..3865
     polyA_signal    3880..3885
     polyA_signal    3900..3905
     polyA_signal    3920..3925
     polyA_signal    3940..3945
     polyA_signal    3960..3965
     polyA_signal    3980..3985
     polyA_signal    4000..4005
     polyA_signal    4020..4025
     polyA_signal    4040..4045
     polyA_signal    4060..4065
     polyA_signal    4080..4085
     polyA_signal    4100..4105
     polyA_signal    4120..4125
     polyA_signal    4140..4145
     polyA_signal    4160..4165
     polyA_signal    4180..4185
     polyA_signal    4200..4205
     polyA_signal    4220..4225
     polyA_signal    4240..4245
     polyA_signal    4260..4265
     polyA_signal    4280..4285
     polyA_signal    4300..4305
     polyA_signal    4320..4325
     polyA_signal    4340..4345
     polyA_signal    4360..4365
     polyA_signal    4380..4385
     polyA_signal    4400..4405
     polyA_signal    4420..4425
     polyA_signal    4440..4445
     polyA_signal    4460..4465
     polyA_signal    4480..4485
     polyA_signal    4500..4505
     polyA_signal    4520..4525
     polyA_signal    4540..4545
     polyA_signal    4560..4565
     polyA_signal    4580..4585
     polyA_signal    4600..4605
     polyA_signal    4620..4625
     polyA_signal    4640..4645
     polyA_signal    4660..4665
     polyA_signal    4680..4685
     polyA_signal    4700..4705
     polyA_signal    4720..4725
     polyA_signal    4740..4745
     polyA_signal    4760..4765
     polyA_signal    4780..4785
     polyA_signal    4800..4805
     polyA_signal    4820..4825
     polyA_signal    4840..4845
     polyA_signal    4860..4865
     polyA_signal    4880..4885
     polyA_signal    4900..4905
     polyA_signal    4920..4925
     polyA_signal    4940..4945
     polyA_signal    4960..4965
     polyA_signal    4980..4985
     polyA_signal    5000..5005
     polyA_signal    5020..5025
     polyA_signal    5040..5045
     polyA_signal    5060..5065
     polyA_signal    5080..5085
     polyA_signal    5100..5105
     polyA_signal    5120..5125
     polyA_signal    5140..5145
     polyA_signal    5160..5165
     polyA_signal    5180..5185
     polyA_signal    5200..5205
     polyA_signal    5220..5225
     polyA_signal    5240..5245
```



Best Local Similarity 97.8%; Pred. No. 1.4e-18;
Matches 544; Conservative 0; Mismatches 9; Indels 3; Gaps 2;



Search completed: January 17, 2003, 06:10:22
Job time : 2821 secs

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 17, 2003, 06:57:47 ; Search time 26 Seconds
(without alignments)
984.005 Million cell updates/sec

Title: US-09-915-873-3
Perfect score: 178
Sequence: 1 accatgcaggtcctgaacac..........gcagcgacctgacctgg[illegible]

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 1816940

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09915873/runat_14012003_171621_13274/[illegible]
-DB=A_Geneseq_101002 -QFMT=fastan -SUFFIX=n2p.rag -MINMATCH=0.1 [illegible]
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09915873_@CGN_1_1_0_@runat_14012003_171621_13274 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqp-embl/[illegible].DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqp-embl/[illegible].DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqp-embl/[illegible].DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqp-embl/[illegible].DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqp-embl/[illegible].DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqp-embl/[illegible].DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqp-embl/[illegible].DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqp-embl/[illegible].DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqp-embl/[illegible].DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqp-embl/[illegible].DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqp-embl/[illegible].DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqp-embl/[illegible].DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqp-embl/[illegible].DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqp-embl/[illegible].DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqp-embl/[illegible].DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqp-embl/[illegible].DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqp-embl/[illegible].DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqp-embl/[illegible].DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqp-embl/[illegible].DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqp-embl/[illegible].DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqp-embl/[illegible].DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqp-embl/[illegible].DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqp-embl/[illegible].DAT:*

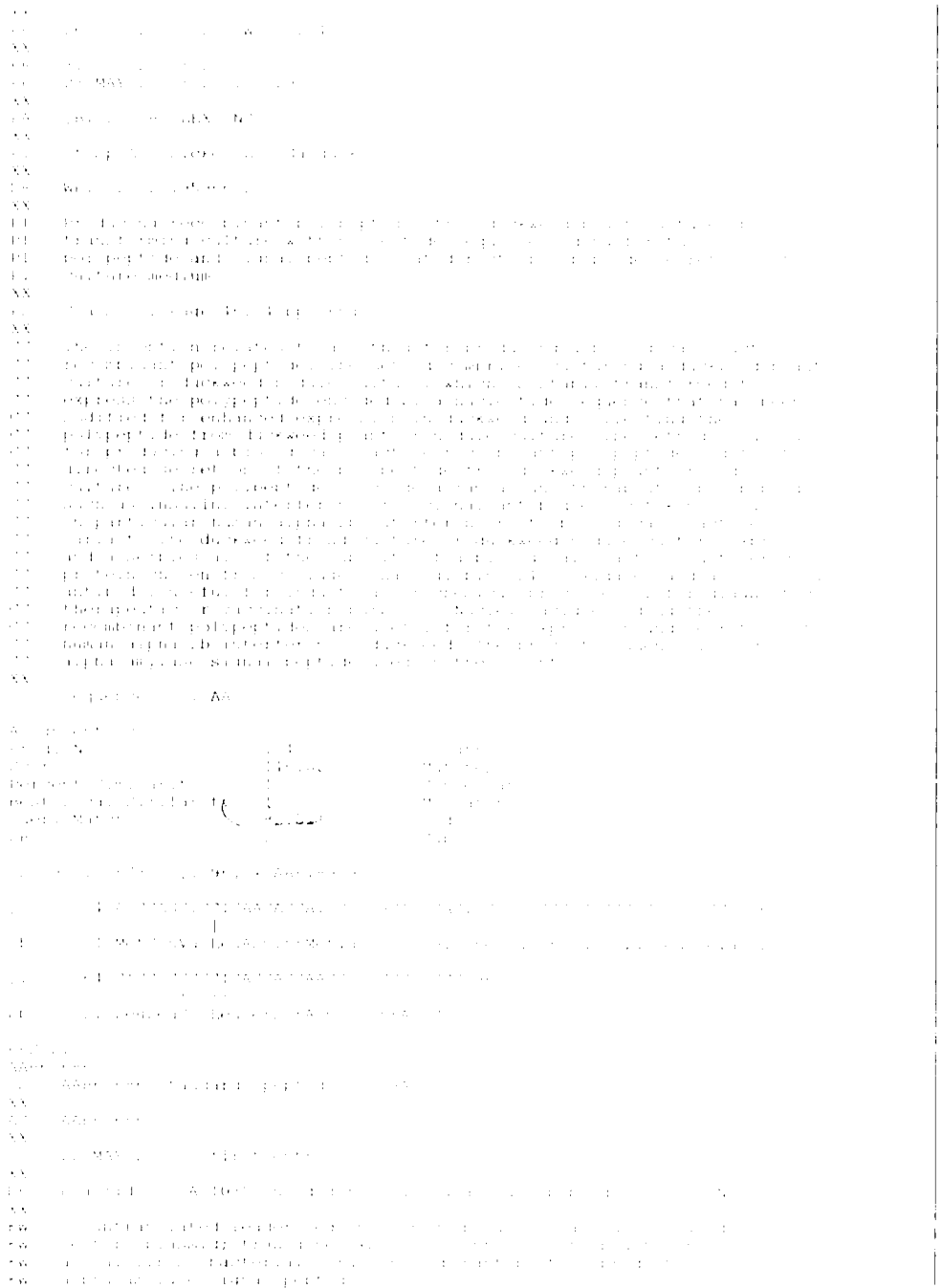
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	146	82.[illegible]	[illegible]	[illegible]	AA[illegible]	[illegible]
2	146	82.[illegible]	[illegible]	22	AAB[illegible]	[illegible]
3	146	82.0	[illegible]	22	AAB[illegible]	[illegible]
4	146	82.[illegible]	[illegible]	22	AAB[illegible]	[illegible]
5	146	82.0	64	22	AAB[illegible]	[illegible]
6	146	82.[illegible]	434	14	AAB[illegible]	[illegible]
7	146	82.0	434	[illegible]	AAW[illegible]	[illegible]
8	146	82.0	434	18	AAW[illegible]	[illegible]
9	146	82.0	434	20	AAY[illegible]	[illegible]
10	146	82.0	434	20	AAW[illegible]	[illegible]
11	146	82.0	434	21	AAY[illegible]	[illegible]
12	132	74.2	[illegible]	[illegible]	AA[illegible]	[illegible]
13	127.5	71.[illegible]	4[illegible]	[illegible]	AA[illegible]	[illegible]
14	117	65.7	[illegible]	[illegible]	AAW[illegible]	[illegible]
15	117	65.7	4[illegible]	[illegible]	AAB[illegible]	[illegible]
16	117	65.7	428	22	AAM[illegible]	[illegible]
17	117	65.7	4[illegible]	2[illegible]	AA[illegible]	[illegible]
18	117	65.7	4[illegible]	1[illegible]	AAB[illegible]	[illegible]
19	[illegible]	[illegible]	[illegible]	[illegible]	AAB[illegible]	[illegible]
20	71.5	4[illegible]	[illegible]	[illegible]	AA[illegible]	[illegible]
21	71.5	[illegible]	[illegible]	[illegible]	AAY[illegible]	[illegible]
22	71.5	40.[illegible]	4[illegible]	21	AAY[illegible]	[illegible]
23	71.5	40.2	48[illegible]	21	AAY[illegible]	[illegible]
24	71.5	40.2	5[illegible]	2[illegible]	AAW[illegible]	BAA [illegible]
25	71.5	40.2	5[illegible]	20	AAW[illegible]	BAA [illegible]
26	71.5	40.2	591	2[illegible]	AAY[illegible]	[illegible]
27	71.5	40.2	591	21	AAY[illegible]	[illegible]
28	71.5	40.2	1000	21	AAY[illegible]	[illegible]
29	71.5	40.2	1000	21	AAY[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	AAB[illegible]	[illegible]
31	66	[illegible]	187	2[illegible]	AAY[illegible]	[illegible]
32	65.5	36.8	2[illegible]	2[illegible]	AAY[illegible]	[illegible]
33	65.5	36.8	114	21	AA[illegible]	[illegible]
34	65	36.5	[illegible]	22	AAB[illegible]	[illegible]
35	64.5	36.2	16[illegible]	22	AAB[illegible]	[illegible]
36	64.5	36.2	17[illegible]	22	AAB[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	AAW[illegible]	[illegible]
38	[illegible]	[illegible]	[illegible]	22	AB[illegible]	[illegible]
39	[illegible]	[illegible]	[illegible]	[illegible]	AB[illegible]	[illegible]
40	[illegible]	[illegible]	[illegible]	[illegible]	AAW[illegible]	[illegible]
41	[illegible]	[illegible]	[illegible]	22	AA[illegible]	[illegible]
42	[illegible]	[illegible]	[illegible]	22	[illegible]	[illegible]
43	[illegible]	[illegible]	[illegible]	22	[illegible]	[illegible]
44	[illegible]	[illegible]	[illegible]	[illegible]	AAB[illegible]	[illegible]
45	[illegible]	[illegible]	[illegible]	22	AA[illegible]	[illegible]

ALIGNMENTS

RESULT 1
AAB[illegible]
ID AAB[illegible] standard; peptide; [illegible] AA.
XX
AC AAB[illegible];
XX
DT [illegible]-MAY-2002 (first entry)
XX
DE Rice alpha-amylase [illegible] peptide.
XX
KW [illegible]
KW [illegible]
KW [illegible]
XX
OS Oryza sativa.
XX
PN WO[illegible]-A[illegible].
XX
PD 07-FEB-2002.



XX
DT 22-MAY-2001 (first entry)
XX
DE pTIUDABP-encoded rice alpha-amylase signal peptide, SEQ ID NO:29.
XX
KW 5' untranslated leader sequence; recombinant viral nucleic acid;
KW vector; plasmid; transgene expression enhancement; plant virus;
KW animal virus; bacterial virus; recombinant protein production; rice
KW alpha-amylase signal peptide.
XX
OS Oryza sativa.
OS Synthetic.
XX
PN WO200107613-A2.
XX
PD 01-FEB-2001.
XX
PF 20-JUL-2000; 2000WO-US20142.
XX
PR 21-JUL-1999; 99US-0359299.
XX
PA (LARG-) LARGE SCALE BIOLOGY CORP.
XX
PI Chapman S, Dawson WO, Donson J, Kumagai MH, Turpen TH;
PI Lindbo JA, Pogue GP, Shivprasad S;
XX
DR WPI; 2001-182793/18.
DR N-PSDB; AAF75925.
XX
PT New recombinant viral nucleic acid for enhancing the production of a
PT protein in a host, comprises a first sequence with a non-native
PT 5'-untranslated sequence, and second sequence linked to the first
PT sequence -
XX
PS Example 6; Fig 6, 47pp, English.
XX
CC The invention relates to a recombinant viral nucleic acid comprising a
CC non-native 5'-untranslated (leader) sequence (S1) upstream of, and
CC operably linked to a second sequence (S2). S2 may be a non-native coding
CC sequence, encoding a product of interest. The presence of the non-native
CC 5'-untranslated leader sequence increases RNA or protein production from
CC S2 relative to the amount produced in the absence of S1. The invention
CC also encompasses a vector and isolated host cell comprising the
CC recombinant viral nucleic acid sequence, and a method of enhancing the
CC production of an RNA or protein in a host using the recombinant viral
CC nucleic acid. The recombinant viral nucleic acid is derived from a
CC plant virus (particularly an RNA plant virus), an animal virus or a
CC bacterial virus, and is used for enhancing the production of RNA or
CC protein (particularly a viral coat protein/non-native protein fusion) in
CC plant, animal or bacterial hosts. The recombinant viral nucleic acid is
CC useful for enhancing the production of RNAs and commercially useful
CC proteins or peptides such as enzymes, antibodies, hormones,
CC pharmaceuticals, vaccines, pigments, and anti-microbial polypeptides.
CC The present sequence represents a rice alpha-amylase signal
CC peptide-containing polypeptide encoded by a plasmid constructed
CC in an exemplification of the invention.
XX
SQ Sequence 33 AA;

Alignment Scores:

Pred. No.:	1.45e-11	Length:	33
Score:	146.00	Matches:	33
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	82.02%	Indels:	0
DB:	22	Gaps:	0

US-09-915-873-3 (1-96) x AAB60788 (1-33)

```
Qy       4 ATGCAGGTCCTGAACACGATGGTCAACAAGCACTTCCTCTCCCTGTCGATCCTCATCGTC 63
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MetGlnValLeuAsnThrMetValAsnLysHisPheLeuSerLeuSerValLeuIleVal 20
```

XX
DT 22-MAY-2001 (first entry)
XX
DE Plasmid REDI encoded rice alpha amylase signal/promoter
XX
KW 5' untranslated leader sequence; recombinant viral nucleic acid;
KW vector; plasmid; transgene expression enhancement; plant virus;
KW animal virus; bacterial virus; recombinant protein production; rice
KW alpha-amylase signal peptide.
XX
OS Chimeric - synthetic
OS Chimeric - Homo sapiens.
XX
PN WO200107613-A2.
XX
PD 01-FEB-2001.
XX
PF 20-JUL-2000; 2000WO-US20142.
XX
PR 21-JUL-1999; 99US-0359299.
XX
PA (LARG-) LARGE SCALE BIOLOGY CORP.
XX
PI Chapman S, Dawson WO, Donson J, Kumagai MH, Turpen TH;
PI Lindbo JA, Pogue GP, Shivprasad S;
XX
DR WPI; 2001-182793/18.
DR N-PSDB; AAF75925.
XX
PT New recombinant viral nucleic acid for enhancing the production of a
PT protein in a host, comprises a first sequence with a non-native
PT 5'-untranslated sequence, and second sequence linked to the first
PT sequence -
XX
PS Example 6; Fig 6, 47pp, English.
XX
CC The invention relates to a recombinant viral nucleic acid comprising a
CC non-native 5'-untranslated (leader) sequence (S1) upstream of, and
CC operably linked to a second sequence (S2). S2 may be a non-native coding
CC sequence, encoding a product of interest. The presence of the non-native
CC 5'-untranslated leader sequence increases RNA or protein production from
CC S2 relative to the amount produced in the absence of S1. The invention
CC also encompasses a vector and isolated host cell comprising the
CC recombinant viral nucleic acid sequence, and a method of enhancing the
CC production of an RNA or protein in a host using the recombinant viral
CC nucleic acid. The recombinant viral nucleic acid is derived from a
CC plant virus (particularly an RNA plant virus), an animal virus or a
CC bacterial virus, and is used for enhancing the production of RNA or
CC protein (particularly a viral coat protein/non-native protein fusion) in
CC plant, animal or bacterial hosts. The recombinant viral nucleic acid is
CC useful for enhancing the production of RNAs and commercially useful
CC proteins or peptides such as enzymes, antibodies, hormones,
CC pharmaceuticals, vaccines, pigments, and anti-microbial polypeptides.
CC The present sequence represents a rice alpha-amylase signal
CC peptide-containing polypeptide encoded by a plasmid constructed
CC in an exemplification of the invention.
XX
SQ Sequence 40 AA;

Alignment Scores:
Pred. No.: 1.45e-11 Length: 33
Score: 146.00 Matches: 33
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

```
SQ   Sequence   434 AA;

Alignment Scores:
Pred. No.:                1.7e-11          Length:          434
Score:                    146.00           Matches:         81
Percent Similarity:       100.00%          Conservative:    0
Best Local Similarity:    100.00%          Mismatches:      0
Query Match:              82.02%           Indels:          0
DB:                       14               Gaps:            0

US-09-915-873-3 (1-96) x AAR32987 (1-434)

Qy       4 ATGCAGGTCCTGAACACGATGGTCAACAACACACTCCTCTCCCTGTCGTCCTCATCGTC 63
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MetGlnValLeuAsnThrMetValAsnLysHisPheLeuSerLeuSerValLeuIleVal 20

Qy      64 CTCCTCGGGCTGAGCAGCAACCTCACGGCGGGC 96
           |||||||||||||||||||||||||||||||||
Db      21 LeuLeuGlyLeuSerSerAsnLeuThrAlaGly 41

RESULT 7
AAW10469
ID   AAW10469 standard; Protein; 434 AA.
XX
AC   AAW10469;
XX
DT   26-APR-1997  (first entry)
XX
DE   Rice alpha-amylase.
XX
KW   Recombinant viral nucleic acid; RNA virus; vector;
KW   tobacco mosaic virus; TMV; rice; alpha-amylase; transgenic plant.
XX
OS   Oryza sativa.
XX
PN   WO9640867-A1.
XX
PD   19-DEC-1996.
XX
PF   06-JUN-1996;   96WO-US09249.
XX
PR   07-JUN-1995;   95US-0483502.
XX
PA   (BIOS-) BIOSOURCE TECHNOLOGIES INC.
XX
PI   Dawson WO,  Donson J,  Garger SJ,  Grantham GL,  Grill LK;
PI   Turpen AM,  Turpen TH;
XX
DR   WPI; 1997-065181/06.
DR   N-PSDB; AAT47095.
XX
PT   Recombinant viral nucleic acid producing, e.g. sole storage in
PT   plants - comprises nucleic acid whose transcription is controlled by
PT   another sequence
XX
PS   Example 4; Page 127-129; 149pp; English.
XX
CC   A DNA sequence (AAT47095) codes for rice alpha-amylase (AAW10469).
CC   Novel viral nucleic acids can be constructed in which the alpha
CC   amylase coding sequence is placed under control of the promoter
CC   of the tobacco mosaic virus coat protein gene, which has been
CC   deleted.  The recombinant viral nucleic acid is capable of cell
CC   replication, encapsidation and systemic spread in infected tobacco
CC   plants, and directs expression of the alpha amylase in plant
CC   tissue.
XX
SQ   Sequence   434 AA;

Alignment Scores:
Pred. No.:                1.7e-11          Length:          434
Score:                    146.00           Matches:         81
Percent Similarity:       100.00%          Conservative:    0
Best Local Similarity:    100.00%          Mismatches:      0
```

```
Query Match:              [illegible]      Indels:          0
DB:                       [illegible]      Gaps:            0

US-09-915-873-3 (1-96) x AAW10469 (1-434)

Qy       4 ATGCAGGTCCTGAACACGATGGTCAACAACACACTCCTCTCCCTGTCGTCCTCATCGTC 63
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MetGlnValLeuAsnThrMetValAsnLysHisPheLeuSerLeuSerValLeuIleVal 20

Qy      64 CTCCTCGGGCTGAGCAGCAACCTCACGGCGGGC 96
           |||||||||||||||||||||||||||||||||
Db      21 LeuLeuGlyLeuSerSerAsnLeuThrAlaGly 41

RESULT 7
AAW10469
ID   AAW10469 standard; Protein; 434 AA.
XX
AC   AAW10469;
XX
DT   26-APR-1997  (first entry)
XX
DE   Rice alpha-amylase.
XX
KW   Recombinant viral nucleic acid; RNA virus; vector;
KW   tobacco mosaic virus; TMV; rice; alpha-amylase; transgenic plant.
XX
OS   Oryza sativa.
XX
PN   WO9640867-A1.
XX
PD   19-DEC-1996.
XX
PF   06-JUN-1996;   96WO-US09249.
XX
PR   07-JUN-1995;   95US-0483502.
XX
PA   (BIOS-) BIOSOURCE TECHNOLOGIES INC.
XX
PI   Dawson WO,  Donson J,  Garger SJ,  Grantham GL,  Grill LK;
PI   Turpen AM,  Turpen TH;
XX
DR   WPI; 1997-065181/06.
DR   N-PSDB; AAT47095.
XX
PT   Recombinant viral nucleic acid producing, e.g. sole storage in
PT   plants - comprises nucleic acid whose transcription is controlled by
PT   another sequence
XX
PS   Example 4; Page 127-129; 149pp; English.
XX
CC   A DNA sequence (AAT47095) codes for rice alpha-amylase (AAW10469).
CC   Novel viral nucleic acids can be constructed in which the alpha
CC   amylase coding sequence is placed under control of the promoter
CC   of the tobacco mosaic virus coat protein gene, which has been
CC   deleted.  The recombinant viral nucleic acid is capable of cell
CC   replication, encapsidation and systemic spread in infected tobacco
CC   plants, and directs expression of the alpha amylase in plant
CC   tissue.
XX
SQ   Sequence   434 AA;

Alignment Scores:
Pred. No.:                1.7e-11          Length:          434
Score:                    146.00           Matches:         81
Percent Similarity:       100.00%          Conservative:    0
Best Local Similarity:    100.00%          Mismatches:      0
```

```
XX
PI   Dawson WO,  Donson J,  Garger SJ,  Grantham GL,  Grill LK;
PI   Turpen AM,  Turpen TH;
XX
DR   WPI; 1999-142035/12.
DR   N-PSDB; AAX03386.
XX
PT   Recombinant plant viral vector  that is capable of systemic
PT   infection in host plant and stable production of heterologous DNA
PT   useful for producing therapeutic proteins for treating e.g. AIDS
XX
PS   Example 4; Columns 51-54; 45pp; English.
XX
CC   The present sequence represents rice alpha-amylase. The nucleic
CC   acid sequence can be expressed in the plant viral constructs of the
CC   invention. The specification describes a recombinant plant viral
CC   nucleic acid derived from a positive (+)-sense RNA plant virus comprising
CC   a native (+)-sense RNA plant virus promoter that is linked to an
CC   expression sequence and a heterologous (+)-sense RNA plant virus
CC   promoter that is linked to an expression sequence. The promoters are
CC   incapable of recombination with each other, and one of the expression
CC   sequences encodes a plant viral coat protein while the other is
CC   optionally a heterologous coding sequence. The plant viral nucleic acid
CC   is capable of systemic infection in a host plant. The viral construct is
CC   useful for the introduction and expression of non viral foreign genes
CC   in plants and the production of e.g. potential AIDS therapeutic drugs.
XX
SQ   Sequence   434 AA;

Alignment Scores:
Pred. No.:                1.7e-11        Length:          434
Score:                    146.00         Matches:         31
Percent Similarity:       100.00%        Conservative:    0
Best Local Similarity:    100.00%        Mismatches:      0
Query Match:              82.02%         Indels:          0
DB:                       20             Gaps:            0

US-09-915-873-3 (1-96) x AAW84383 (1-434)

Qy       4 ATGCAGGTCCTGAACACGATGGTCAACAACCACTTCCTCCTCTCCTCCAACTCTCTCGCC 63
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MetGlnValLeuAsnThrMetValAsnLysHisPheLeuSerLeuSerSerAsnSerLeuAla 60

Qy      64 CTCCTCGGGCTGAGCAGCAACCTCACCGCCGGC 96
           |||||||||||||||||||||||||||||||||
Db      21 LeuLeuGlyLeuSerSerAsnLeuThrAlaGly 53

RESULT 11
AAY87792
ID   AAY87792 standard; Protein; 434 AA.
XX
AC   AAY87792;
XX
DT   24-AUG-2000  (first entry)
XX
DE   Rice alpha-amylase protein.
XX
KW   Animal RNA virus; viral coat protein; plant; male sterility;
KW   interleukin, EPO, erythropoietin; CSF; colony stimulating factor;
KW   Factor VIII; hGH; human growth hormone; melanin; insulin; vaccine;
KW   stereo specific catalysis; alpha-amylase; rice.
XX
OS   Oryza sativa.
XX
PN   US6054566-A.
XX
PD   25-APR-2000.
XX
PF   07-JUN-1995;   95US-0484341.
XX
PR   17-FEB-1989;   89US-0310881.
PR   08-JUN-1989;   89US-0364138.
PR   26-FEB-1988;   88US-0160766.
```

XX
AC AAW59838;
XX
DT 20-NOV-1998 (first entry)
XX
DE RAmy1A signal peptide amino acid sequence.
XX
KW RAmy1A signal peptide; protein expression; monocotyledon plant cell;
KW glycosylated alpha 1 antitrypsin; AAT; glycosylated antithrombin III;
KW ATIII; human serum albumin; HSA; subtilisin BPN'; treatment; emphysema;
KW antithrombotic; blood replacement.
XX
OS Unidentified.
XX
PN WO9836085-A1.
XX
PD 20-AUG-1998.
XX
PF 13-FEB-1998; 98WO-US003068.
XX
PR 13-FEB-1997; 97US-0038170.
PR 13-FEB-1997; 97US-0037991.
PR 13-FEB-1997; 97US-0038168.
PR 13-FEB-1997; 97US-0038169.
XX
PA (PHYT-) APPLIED PHYTOLOGICS INC.
XX
PI Rodriguez RL, Sutliff TD;
XX
DR WPI; 1998-467179/40.
XX
PT Expressing mature, glycosylated proteins in monocotyledonous plant
PT cells - from chimeric gene including signal peptide sequence,
PT specifically therapeutic agents and industrial enzymes
XX
PS Claim 6; Page 27; 54pp; English.
XX
CC The present sequence represents the RAmy1A signal peptide sequence. It
CC is used in the course of the invention. The specification describes a
CC method for producing mature heterologous protein in monocotyledonous
CC plant cells. The method comprises transforming the cells with a chimeric
CC gene comprising a monocotyledon transcription regulator, inducible either
CC during seed maturation or by adding/removing a cell culture, DNA
CC encoding the heterologous protein, and DNA encoding a signal peptide,
CC with the signal peptide causing secretion of the protein from the cell.
CC Proteins expressed in this manner include mature glycosylated alpha
CC 1-antitrypsin (AAT) with a glycosylation pattern that significantly
CC increases its serum half-life, mature glycosylated antithrombin III
CC (ATIII), mature human serum albumin (HSA) having the native folding
CC pattern as shown by bilirubin-binding characteristics, or mature active
CC subtilisin BPN'. These proteins are useful therapeutically (e.g. AAT for
CC treating emphysema, ATIII as antithrombotic and HSA as blood replacement)
CC or as industrial enzymes (BPN' is used in detergents).
XX
SQ Sequence 25 AA;

Alignment Scores:

Pred. No.:	1.31e-07	Length:	25
Score:	117.00	Matches:	25
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	65.73%	Indels:	0
DB:	19	Gaps:	0

US-09-915-873-3 (1-96) x AAW59838 (1-25)

```
Qy      22 ATGGTCAACAAGCACTTCCTCTCCCTCGTCGTCCTCGTCCTCGTCTCCTCCTCCTCC 81
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MetValAsnLysHisPheLeuSerLeuValValLeuValLeuValSerSerSerSer 20

Qy      82 AACCTCACCGCCGGC 96
           |||||||||||||||
Db      21 AsnLeuThrAlaGly 25
```

RESULT 2
AAB12798
ID AAB12798 standard; protein; 25 AA.
XX
AC AAB12798;
XX
DT 23-NOV-1998 (first entry)
XX
DE Rice alpha Amy7 promoter sequence.
XX
KW Rice; alpha Amy7; alpha Amy7; alpha Amy; alpha Amy;
KW gene expression; promoter; cereal; alpha-amylase; transgenic plant.
XX
OS Oryza sativa.
XX
PN EP-842999-A.
XX
PD 20-MAY-1998.
XX
PF 14-NOV-1997; 97EP-0309135.
XX
PR 15-NOV-1996; 96JP-0305116.
PR 21-NOV-1996; 96JP-0310109.
XX
PA (NASH/) NAT FOOD RES INST.
XX
DR WPI; 1998-281741/24.
DR N-PSDB; AAB12797.
XX
PT A gene expression system comprising the promoter region of
PT alpha-amylase gene, and a DNA sequence coding a desired protein
PT product in a plant cell
XX
PS Example 1; Page 27; 44pp; Japanese.
XX
CC The present invention describes a gene expression system comprising the
CC promoter region of an alpha amylase gene. Also described is a method for
CC preparing a transgenic rice plant comprising (1) introducing a DNA
CC into an immature embryo of a rice plant; (2) simultaneously culturing
CC the embryo with a suspended culture of a dicotyledon derived for
CC transformation step; (3) at which the transformed embryo is cultured in
CC selective medium containing plant growth hormones and (4) a callus is
CC the callus to roots and sprouts in a redifferentiation medium. The DNA is
CC be used for mass production of a desired protein in plant cells. The
CC cell. The present sequence represents the rice alpha Amy7 promoter
CC which is used in an example to illustrate the present invention.
XX
SQ Sequence 25 AA;

Alignment Scores:

Pred. No.:	1.31e-07	Length:	25
Score:	117.00	Matches:	25
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	65.73%	Indels:	0
DB:	21	Gaps:	0

US-09-915-873-3 (1-96) x AAB12798 (1-25)

```
Qy      22 ATGGTCAACAAGCACTTCCTCTCCCTCGTCGTCCTCGTCCTCGTCTCCTCCTCCTCC 81
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MetValAsnLysHisPheLeuSerLeuValValLeuValLeuValSerSerSerSer 20

Qy      82 AACCTCACCGCCGGC 96
           |||||||||||||||
Db      21 AsnLeuThrAlaGly 25
```

Search completed: January 17, 2003, 09:14:21
Job time : 1.31 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2003, 05:07:40 ; Search time 27.5 seconds
(without alignments)
4569.989 Million cell updates/sec

Title: US-09-915-873-1
Perfect score: 554
Sequence: 1 gatcaagtgaaaaatgcac.....ctatcaatgtgatctcga 554

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/Geneseq/geneseq-embl/NA_COMB.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseq-embl/NA1.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseq-embl/NA2.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseq-embl/NA3.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseq-embl/NA4.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseq-embl/NA5.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseq-embl/NA6.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseq-embl/NA7.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseq-embl/NA8.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseq-embl/NA9.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseq-embl/NA10.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseq-embl/NA11.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseq-embl/NA12.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseq-embl/NA13.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseq-embl/NA14.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseq-embl/NA15.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseq-embl/NA16.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseq-embl/NA17.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseq-embl/NA18.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseq-embl/NA19.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseq-embl/NA20.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseq-embl/NA21.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseq-embl/NA22.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseq-embl/NA23.DAT:*

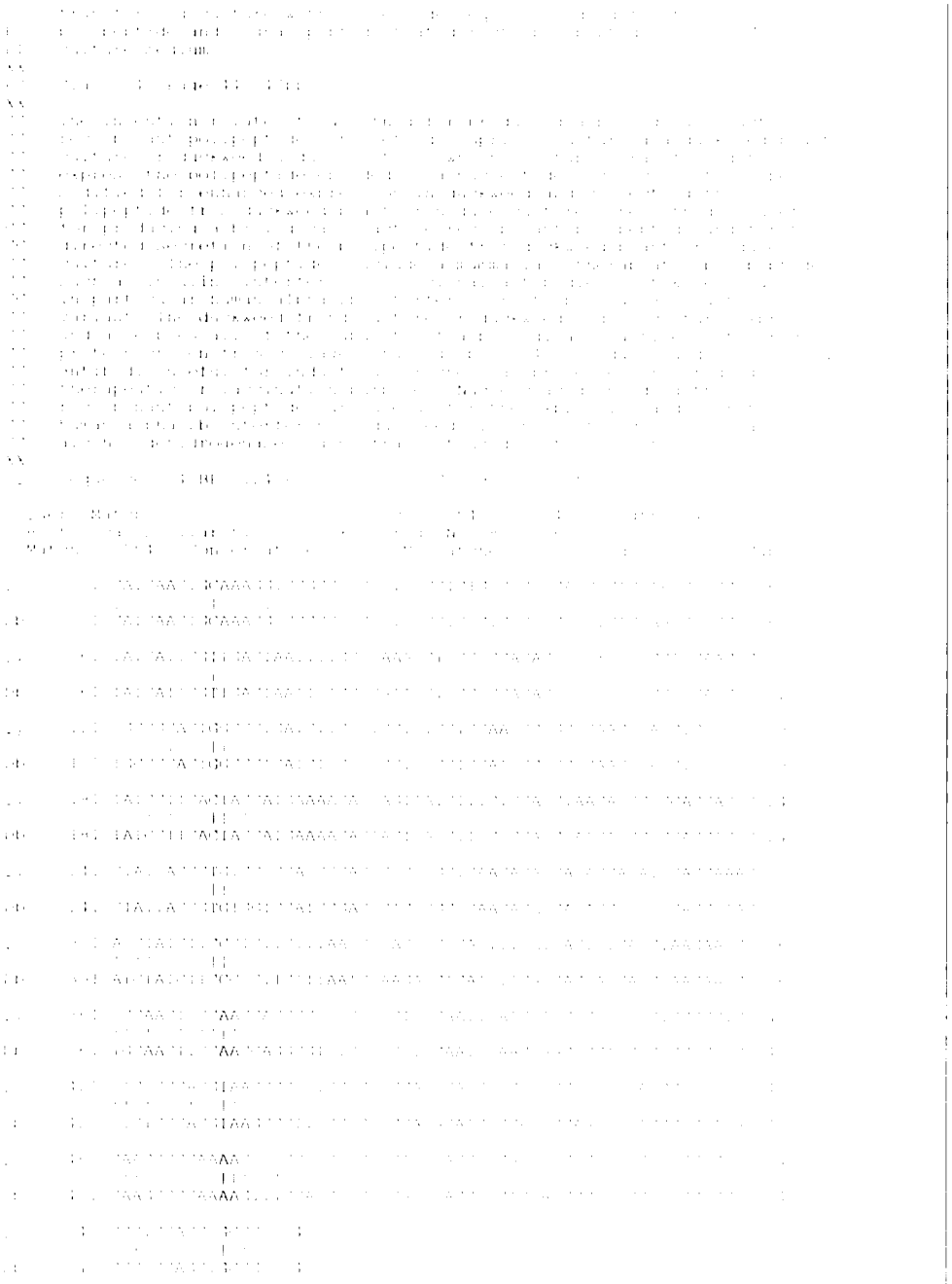
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

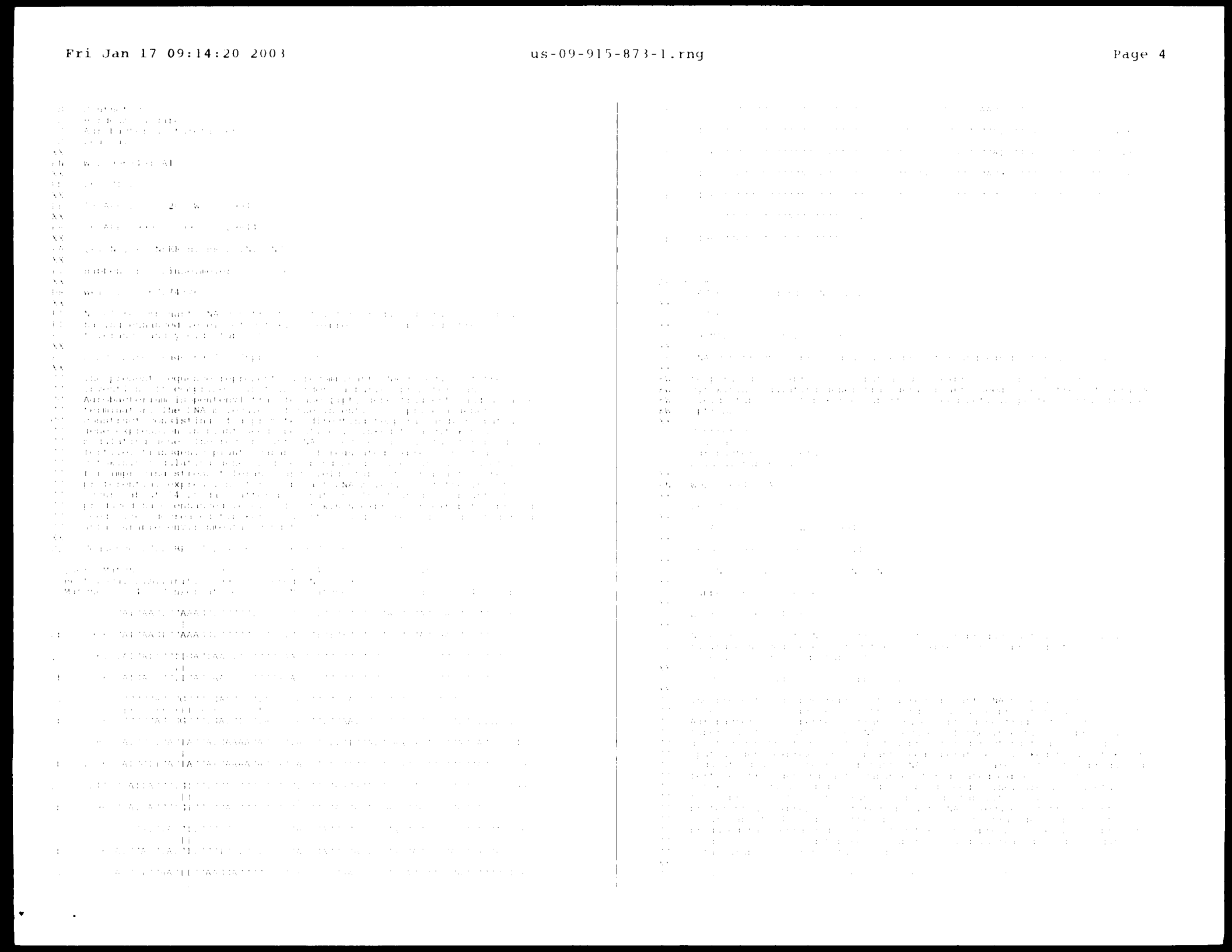
Result No.	Score	Query Match	Length	DB	ID	Description
1	554	100.0	554	24	AAI38345	Maize alcohol deh
2	541.4	97.7	1086	19	AAV54017	Coding region of
3	541.4	97.7	1086	22	AAC68665	Azrobacterium tu
4	540	97.5	2722	21	AAC86507	DNA cognistin 1
5	540	97.5	2722	21	AAC86508	DNA construct seq
6	540	97.5	5033	15	AAQ74695	pPHI1406 vector co
7	532	96.0	1343	21	AAC87199	Rice ESTS promote
8	532	96.0	1343	21	AAC88404	Rice genomic 5.1
9	532	96.0	1343	21	AAC89428	Rice genomic 5.11

ALIGNMENTS

RESULT 1
AAI38345
ID AAI38345 standard; DNA; [illegible]
XX
AC AAI38345;
XX
DT [illegible] (first entry)
XX
DE Maize alcohol dehydrogenase 1 gene fragment.
XX
KW Maize; monocot plant; alpha alcohol dehydrogenase; enzyme; [illegible]
XX
OS Zea mays.
XX
RN [illegible]
XX
RC [illegible]
XX
RP [illegible]
XX
RP [illegible]
XX
RA [illegible]
XX
RT Stemp A., Dickey L., Gustafson J.;
XX
RL [illegible]
XX
DR WPI; [illegible].
XX
RT Producing recombinant polypeptide [illegible]



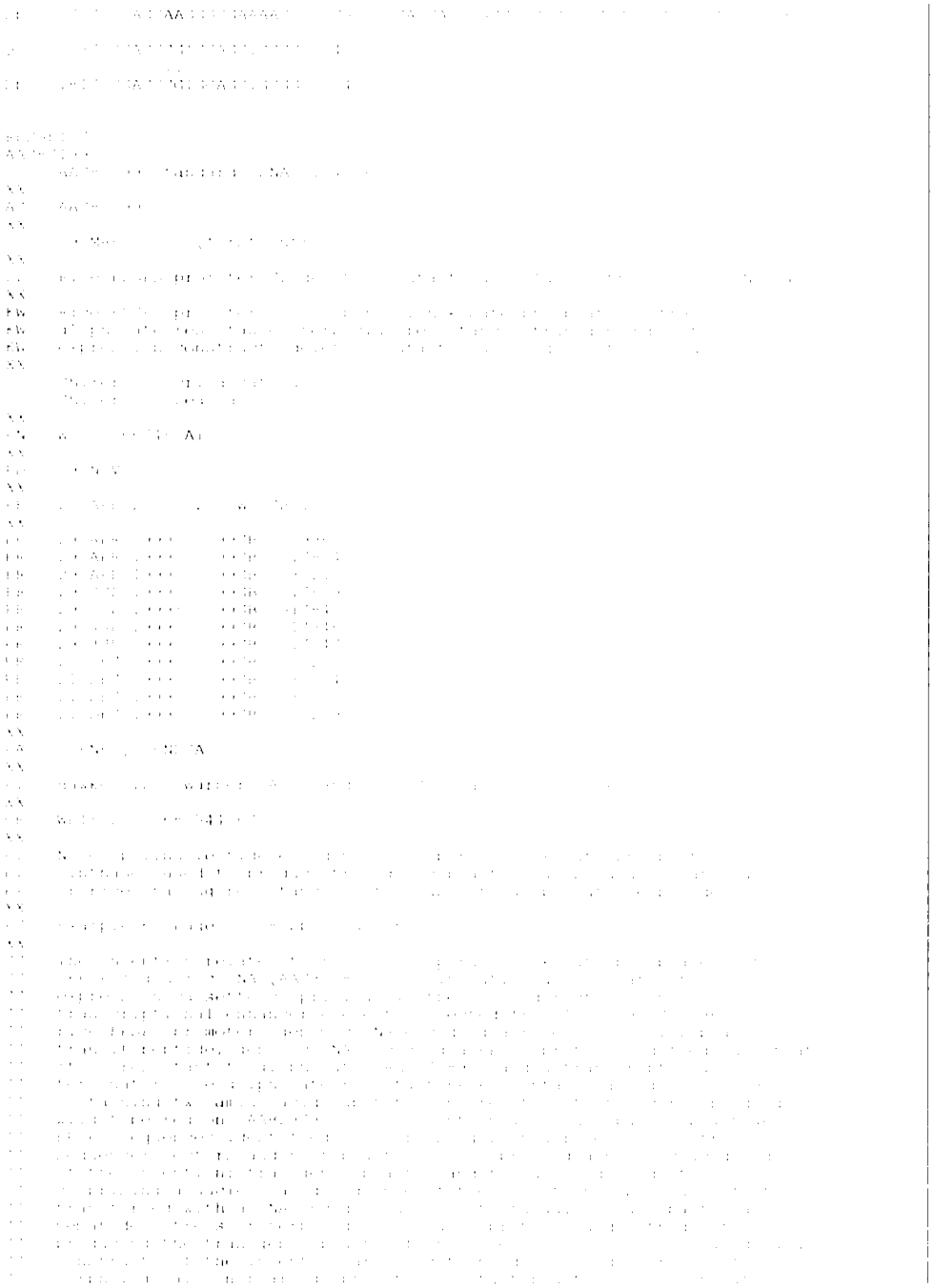
```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        349 GTGGGCTGATCTTGTATGCTATCCTGCAATCGTCCTCAACTTATCTCTTTATATCCTT 408

Qy        188 CACTACCATGAAAAGACTAGTAATCTTTCTCGATGTAACATCGTCCAGCACTGCTATTAC 247
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        409 CACTACCATGAAAAGACTAGTAATCTTTCTCGATGTAACATCGTCCAGCACTGCTATTAC 468

Qy        248 CGTCTCGTCCATCCGACAGTCTGGCTGAACACATCATACGATATTGAGCAAAGATCTATC 307
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        469 CGTGTGGTCCATCCGACAGTCTGGCTGAACACATCATACGATATTGAGCAAAGATCTATC 528

Qy        308 TTCCCTGTTCTTTAATGAAAGACGTCATTTTCATCAGTATGATCTAAGAATGTTGCAACT 367
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        529 TTCCCTGTTCTTTAATGAAAGACGTCATTTTCATCAGTATGATCTAAGAATGTTGCAACT 588

Qy        368 TGCAAGGAGGCGTTTCTTTCTTTGAATTTAACTAACTCGTTGAGTGGCCCTGTTTCTCGG 427
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        589 TGCAAGGAGGCGTTTCTTTCTTTGAATTTAACTAACTCGTTGAGTGGCCCTGTTTCTCGG 648

Qy        428 ACGTAAGGCCTTTGCTGCTCCACACATGTCCATTCGAATTTTACCGTGTTTAGCAAGGGC 487
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        649 ACGTAAGGCCTTTGCTGCTCCACACATGTCCATTCGAATTTTACCGTGTTTAGCAAGGGC 708

Qy        488 GAAAAGTTTGCATCTTGATGATTTAGCTTGACTATGCGATTGCTTTCCTGGACCCGTGCA 547
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        709 GAAAAGTTTGCATCTTGATGATTTAGCTTGACTATGCGATTGCTTTCCTGGACCCGTGCA 768

Qy        548 GCT 550
              |||
Db        769 GCT 771


RESULT 3
AAC68665
ID   AAC68665 standard; DNA; 1086 BP.
XX
AC   AAC68665;
XX
DT   21-FEB-2001  (first entry)
XX
DE   Agrobacterium tumefaciens pGVS71 T-DNA vector bar gene.
XX
KW   Agrobacterium tumefaciens; T-DNA; transgenic plant; DNA transfer;
KW   bar gene; adhI intron; ds.
XX
OS   Chimeric - Agrobacterium tumefaciens.
OS   Chimeric - Zea mays.
XX
PN   US6140553-A.
XX
PD   31-OCT-2000.
XX
PF   20-FEB-1998;   98US-0026673.
XX
PR   20-FEB-1997;   97US-0135507.
XX
PA   (PLBZ ) PLANT GENETIC SYSTEMS NV.
XX
PI   D'Halluin K;
XX
DR   WPI; 2001-006346/01.
XX
PT   Improved method for integrating a DNA fragment into the genome of a
PT   monocot plant cell, useful for increasing the transformation efficiency
PT   of DNA transfer, involves pre-treating the cells in a medium containing
PT   plant phenolics -
XX
PS   Disclosure; Column 19-22; 13pp; English.
XX
CC   The present sequence is the coding region of the bar gene in the T-DNA
CC   vector pGVS71. The gene contains the adhI intron from maize. The vector
CC   was introduced into Agrobacterium tumefaciens in a method for integrating
CC   a DNA fragment into the genome of a cell of a monocotyledonous plant. The
CC   method comprises incubating, prior to contacting with the DNA fragment, a
CC   culture of untransformed [illegible] plant cells [illegible]
CC   containing a plant phenolic compound. The method is useful for improved
CC   Agrobacterium-mediated DNA transfer to plant cells (particularly [illegible]
CC   cells), or plant cell cultures (particularly of monocotyledonous plants)
CC   that have been pre-treated with plant phenolics. The plant phenolics
CC   stimulate cell division and enhance competence for integration of foreign
CC   DNA. The method provides a rapid, efficient and reproducible method for
CC   increasing the transformation efficiency of DNA transfer, particularly
CC   Agrobacterium-mediated DNA transfer.
XX
SQ   Sequence 1086 BP; [illegible]

  Query Match             [illegible]
  Best Local Similarity   [illegible]
  Matches  542;  Conservative  [illegible]  Mismatches  [illegible]  Indels  [illegible]

Qy        [illegible]
Db        [illegible]

Qy        188 CACTACCATGAAAAGACTAGTAATC[illegible]
Db        409 CACTACCATGAAAAGACTAGTAATC[illegible]

Qy        248 [illegible]
Db        469 [illegible]

Qy        308 [illegible]
Db        529 [illegible]

Qy        368 [illegible]
Db        589 [illegible]

Qy        428 [illegible]
Db        649 [illegible]

Qy        488 [illegible]
Db        709 [illegible]

Qy        548 GCT 550
              |||
Db        769 GCT 771


RESULT 4
AA[illegible]
ID   AA[illegible] standard; DNA; [illegible]
XX
AC   AA[illegible];
XX
DT   [illegible]  (first entry)
XX
DE   DNA comprising [illegible]
XX
KW   temporal gene expression; [illegible]
KW   cytokinin [illegible]
KW   yield stability; [illegible]
KW   ipt; ss.
XX
```

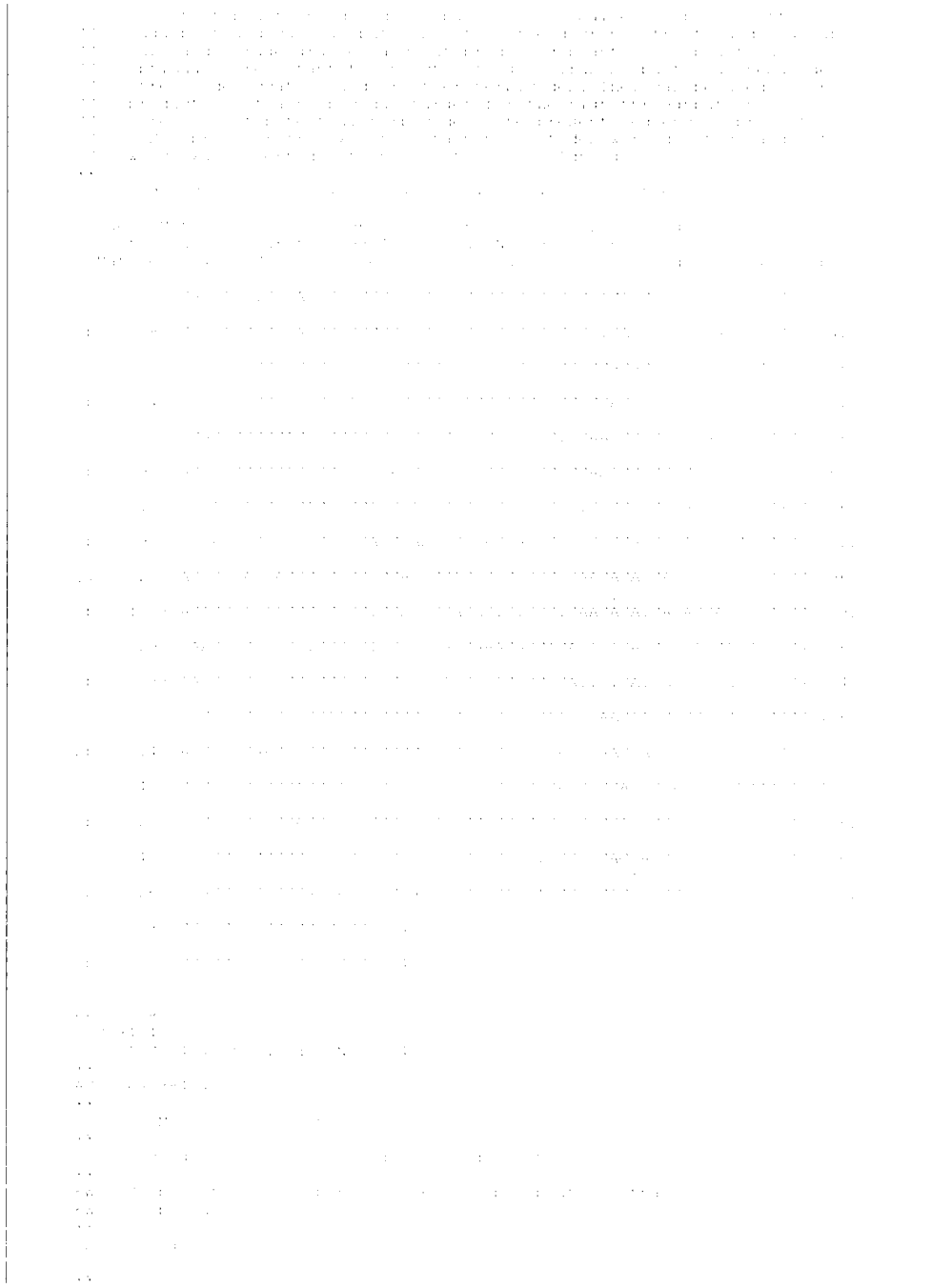


```
  Query Match             97.5%;  Score 540;  DB 21;  Length 2722;
  Best Local Similarity   99.3%;  Pred. No. 4.9e-158;
  Matches  554;  Conservative    0;  Mismatches    0;  Indels    4;  Gaps    1;

Qy       1 GATCAAGTGCAAAGGTCCGGCTTGTTTCTCTCTGTCCTTGATCTGATAATAATTTTGTTT 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     956 GATCAAGTGCAAAGGTCCGGCTTGTTTCTCCTCTGTCCTTGATCTGATAATAATTTTGTTT 1015

Qy      61 TATGATTCGTTGAGTAATTTTGGGGAAAGCTTCGTCCACAGTTTTTTTTTTCGATGAATAT 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1016 TATGATTCGTTGAGTAATTTTGGGGAAAGCTTCGTCCACAGTTTTTTTTTCGATGAATAT 1075

Qy     121 TGCCGCAGTGGCGCTGATCTTGTATGCTATCCTGCAATGCTGCTGAATTAATCTTTTTA 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1076 TGCCGCAGTGGCGCTGATCTTGTATGCTATCCTGCAATGCTGCTGAATTAATCTTTTTA 1135

Qy     181 TATCCTTCACTACCATGAAAAGACTAGTAATCTTTCTCGATGTAACATCGTCCACGACTG 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1136 TATCCTTCACTACCATGAAAAGACTAGTAATCTTTCTCGATGTAACATCGTCCACGACTG 1195

Qy     241 CTATTACCGTGTGGTCCATCCGACAGTCTGGCTGAACACATCATATCGATATGAATAAAG 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1196 CTATTACCGTGTGGTCCATCCGACAGTCTGGCTGAACACATCATATCGATATGAATAAAG 1255

Qy     300 ---GATCTATCTTCCTCTTCTTTAATGAAAGACGTCATTTCTATCATGAATCATCTAAATA 356
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1256 ATCGATCTATCTTCCTCTTCTTTAATGAAAGACGTCATTTCTATCATGAATCATCTAAATA 1315

Qy     357 ATGTTGCAACTTGCAAGGAGCGCTTTCTTTCTTTGAATTTAACTAACTCGTTGCAGTGGCT 416
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1316 ATGTTGCAACTTGCAAGGAGCGCTTTCTTTCTTTGAATTTAACTAACTCGTTGCAGTGGCT 1375

Qy     417 CTGTTTCTCGGACGTAAGCCCTTTCTCTCCACACATGTCCATCCGAATTTTACTTTTCT 476
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1376 CTGTTTCTCGGACGTAAGCCCTTTCTCTCCACACATGTCCATCCGAATTTTACTTTTCT 1435

Qy     477 TTAGCAAGGGCGAAAAGTTTGCATCTTGATGATTTAGCTTGACTATGCATTGCTTCCT 536
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1436 TTAGCAAGGGCGAAAAGTTTGCATCTTGATGATTTAGCTTGACTATGCATTGCTTCCT 1495

Qy     537 GGACCCGTGCAGCTGCGG 554
           ||||||||||||||||||
Db    1496 GGACCCGTGCAGCTGCGG 1513


RESULT 6
AAQ74695
ID   AAQ74695 standard; DNA; 5033 BP.
XX
AC   AAQ74695;
XX
DT   09-MAY 1995  (first entry)
XX
DE   pPHI1406 vector containing the MCMV coat protein 3' genes.
XX
KW   MCMV; maize chlorotic dwarf virus; coat protein; plasmid vector;
KW   plant resistance; viral infection resistance; maize plant;
KW   maize dwarf mosaic virus; MDMV; ds.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   CDS             119..672
FT                   /*tag= a
FT                   /note= "Zea mays alcohol dehydrogenase 1, intron 1
FT                   fragment"
XX
PN   WO9421796-A.
XX
PD   29-SEP-1994.
XX
PF   22-MAR-1994;   94WO-US03028.
XX
```

```
PR   24-MAR-1993;  93US-0036928.
XX
PA   (PIOS ) PIONEER HI BRED INT INC
PA   (GENA ) US SEC OF AGRIC.
XX
PI   Mcmullen MD;  Hoch DA;  Louwerse JD;
XX
DR   WPI; 1994-312726/38.
XX
P2   DNA encoding maize chlorotic dwarf virus coat protein - for transformation
P2   of maize plants with resistance to the virus and to related viral
P2   infections
XX
PS   Claim 2; Page 14; 41pp; English.
XX
CC   This sequence shows the nucleotide sequence for the pPHI1406 vector
CC   pPHI1406. The plasmid contains a gene for the coat protein of MCMV (maize
CC   dwarf virus) coat protein expressed under control of the 35S promoter
CC   maize virus 35S promoters isolated from the MDMV strain. It also contains
CC   and a polyA signal sequence obtained from the potato proteinase
CC   inhibitor II (Pin II) gene that exhibits enhancer-like activity. The
CC   chimeric gene also includes a 5' UTR sequence which is a fragment
CC   region of tobacco mosaic virus that functions as a translation
CC   enhancer and a Zea mays alcohol dehydrogenase 1 intron 1 fragment
CC   spanning nucleotides 1266-2176 (SEQ ID NO:2) with Pin II 3' UTR
CC   which has been shown to function as an efficient 3' poly-adenylation
CC   sequence. The DNA is used for imparting resistance to MCMV
CC   viruses to which MCMV infection resistance provides
CC   cross-resistance, including maize dwarf mosaic virus strain A, or
CC   all of the three coat protein genes of MCMV can be used for
CC   protection of plants. MCMV has a single-stranded RNA
CC   (c.AA2745).
XX
SQ   Sequence 5033 BP; 1547 A; 1277 C; 1277 G; 932 T; 0 other;
```

```
  Query Match             97.5%;  Score 540;  DB 21;  Length 2722;
  Best Local Similarity   99.3%;  Pred. No. 4.9e-158;
  Matches  554;  Conservative    0;  Mismatches    0;  Indels    4;  Gaps    1;

Qy       1 GATCAAGTGCAAAGGTCCGGCTTGTTTCTCTCTGTCCTTGATCTGATAATAATTTTGTTT 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1277 GATCAAGTGCAAAGGTCCGGCTTGTTTCTCCTCTGTCCTTGATCTGATAATAATTTTGTTT 1336

Qy      61 TATGATTCGTTGAGTAATTTTGGGGAAAGCTTCGTCCACAGTTTTTTTTTCGATGAATAT 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1337 TATGATTCGTTGAGTAATTTTGGGGAAAGCTTCGTCCACAGTTTTTTTTTCGATGAATAT 1396

Qy     121 TGCCGCAGTGGCGCTGATCTTGTATGCTATCCTGCAATGCTGCTGAATTAATCTTTTTA 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1397 TGCCGCAGTGGCGCTGATCTTGTATGCTATCCTGCAATGCTGCTGAATTAATCTTTTTA 1456

Qy     181 TATCCTTCACTACCATGAAAAGACTAGTAATCTTTCTCGATGTAACATCGTCCACGACTG 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1457 TATCCTTCACTACCATGAAAAGACTAGTAATCTTTCTCGATGTAACATCGTCCACGACTG 1516

Qy     241 CTATTACCGTGTGGTCCATCCGACAGTCTGGCTGAACACATCATATCGATATGAATAAAG 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1517 CTATTACCGTGTGGTCCATCCGACAGTCTGGCTGAACACATCATATCGATATGAATAAAG 1576

Qy     300 ---GATCTATCTTCCTCTTCTTTAATGAAAGACGTCATTTCTATCATGAATCATCTAAATA 356
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1577 ATCGATCTATCTTCCTCTTCTTTAATGAAAGACGTCATTTCTATCATGAATCATCTAAATA 1636

Qy     357 ATGTTGCAACTTGCAAGGAGCGCTTTCTTTCTTTGAATTTAACTAACTCGTTGCAGTGGCT 416
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1637 ATGTTGCAACTTGCAAGGAGCGCTTTCTTTCTTTGAATTTAACTAACTCGTTGCAGTGGCT 1696

Qy     417 CTGTTTCTCGGACGTAAGCCCTTTCTCTCCACACATGTCCATCCGAATTTTACTTTTCT 476
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1697 CTGTTTCTCGGACGTAAGCCCTTTCTCTCCACACATGTCCATCCGAATTTTACTTTTCT 1756

Qy     477 TTAGCAAGGGCGAAAAGTTTGCATCTTGATGATTTAGCTTGACTATGCATTGCTTCCT 536
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1757 TTAGCAAGGGCGAAAAGTTTGCATCTTGATGATTTAGCTTGACTATGCATTGCTTCCT 1816

Qy     537 GGACCCGTGCAGCTGCGG 554
           ||||||||||||||||||
Db    1817 GGACCCGTGCAGCTGCGG 1834
```





PN WO200066746-A1.
XX
PD 09-NOV-2000.
XX
PF 20-APR-2000; 2000WO-GB01559.
XX
PR 29-APR-1999; 99GB-0009971.
PR 29-APR-1999; 99GB-0009972.
PR 29-JUL-1999; 99GB-0017837.
PR 29-JUL-1999; 99GB-0017842.
PR 21-DEC-1999; 99GB-0030190.
PR 21-DEC-1999; 99GB-0030206.
PR 21-DEC-1999; 99GB-0030214.
PR 21-DEC-1999; 99GB-0030216.
XX
PA (ZENE) ZENECA LTD.
XX
PI Hawkes TR, Warner SAJ, Andrews CJ, Barlow S, Pickerill AI;
XX
DR WPI; 2000-670763/66.
XX
PT Novel polynucleotide encoding the rice 5-enolpyruvylshikimate phosphate
PT synthase, used to produce glyphosate tolerant or resistant plants
XX
PS Disclosure; Page 56; 85pp; English.
XX
CC The present invention relates to a glyphosate resistant rice
CC 5-enolpyruvylshikimate phosphate synthase (EPSPS) gene. This gene can
CC be used to produce plant tissue and/or morphologically normal fertile
CC whole plants which are tolerant or resistant to glyphosate herbicides,
CC and in the production of a herbicidal target for the high throughput
CC in vitro screening of potential herbicides.
XX
SQ Sequence 1343 BP; 349 A; 344 C; 245 G; 405 T; 0 other;

Query Match 96.0%; Score 542; DB 21; Length 1343;
Best Local Similarity 98.4%; Pred. No. 9.2e-156;
Matches 549; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

Qy 1 GATCAAGTGCAAAGGTCCGCCTTGTTTCTCCTCTGTCTCTTGATCTGATTAATCTTTGTTT 60
Db 783 GATCAAGTGCAAAGGTCCGCCTTGTTTCTCCTCTGTCTCTTGATCTGATTAATCTTTGTTT 842

Qy 61 TATGATTCGTTGAGTAATTTTGGGGCAA----AGCTTCGTCCACACATTTTTTTTTCTCAATCA 116
Db 843 TATGATTCGTTGAGTAATTTTGGGGCAAAGCTAGCTTCGTCCACACATTTTTTTTTCTCAATCA 902

Qy 117 ACAGTGCCGCAGTGGCGCTGATCTTGTATGCTATCCTGCAATCTGGGTGAACTTATGTCT 176
Db 903 ACAGTGCCGCAGTGGCGCTGATCTTGTATGCTATCCTGCAATCTGCCTGAACTTATTTCT 962

Qy 177 TTTATATCCTTCACTACCATGAAAAGACTAGTAATCTTTCTCGATGTAACATCTCCAGC 236
Db 963 TTTATATCCTTCACTCCCATGAAAAGCCTAGTAATCTTTCTCGATGTAACATCGTCCAGC 1022

Qy 237 ACTGCTATTACCGTGTGGTCCATCCGACAGTCTGGCTGAACACATCATACGATATTGAGC 296
Db 1023 ACTGCTATTACCGTGTGGTCCATCCGACAGTCTGGCTGAACACATCATACGATATTGAGC 1082

Qy 297 AAAGATCTATCTTCCCTGTTCTTTAATGAAAGACGTCATTTTCATCAGTATGATCTAAGA 356
Db 1083 AAAGATCTATCGTCGCTGTTCTTTAATGAAAGACGTCATTTTCATCAGTATGATCTAAGA 1142

Qy 357 ATGTTGCAACTTGCAAGGAGGCGTTCTTCTTTGAATTTAACTAACTCGTTGACTCTT 416
Db 1143 ATGTTGCAACTTGCAAGGAGGCGTTCTTCTTTGAATTTAACTAACTCGTTGACTCTT 1202

Qy 417 CTGTTTCTCGGACGTAAGGCCTTTGCTGCTCCACACATGCCATGGAATTTTATCCTC 476
Db 1203 CTGTTTCTCGGACGTAAGGCCTTTGCTGCTCCACACATGCCATGGAATTTTATCCTC 1262

Qy 477 TTAGCAAGGGCGAAAAGTTTGCATCTTGATGATTAGCTCACTATCGATTGCTTCCT 536
Db 1263 TTAGCAAGGGCGAAAAGTTTGCATCTTGATGATTAGCTCACTATCGATTGCTTCCT 1322

Qy 537 GCATCCTCTGCCT 549
Db 1323 GCATCCTCTGCCT 1335

RESULT 9
AAC84628
ID AAC84628 standard; DNA; 1343 BP.
XX
AC AAC84628;
XX
DT 29-MAR-2001 (first entry)
XX
DE Rice 5-enolpyruvylshikimate-3-phosphate synthase gene.
XX
KW 5-enolpyruvylshikimate-3-phosphate synthase; EPSPS;
KW herbicide resistance; herbicide tolerance; rice.
XX
OS Oryza sp.
XX
PN WO200066747-A1.
XX
PD 09-NOV-2000.
XX
PF 20-APR-2000; 2000WO-GB01559.
XX
PR 29-APR-1999; 99GB-0009971.
PR 29-APR-1999; 99GB-0009972.
PR 29-APR-1999; 99GB-0009973.
PR 29-APR-1999; 99GB-0009974.
PR 29-APR-1999; 99GB-0009975.
PR 29-JUL-1999; 99GB-0017837.
PR 29-JUL-1999; 99GB-0017842.
PR 21-DEC-1999; 99GB-0030190.
PR 21-DEC-1999; 99GB-0030206.
PR 21-DEC-1999; 99GB-0030214.
PR 21-DEC-1999; 99GB-0030216.
XX
PA (ZENE) ZENECA LTD.
XX
PI Hawkes TR, Warner SAJ, Andrews CJ, Barlow S, Pickerill AI;
XX
DR WPI; 2000-670764/66.
XX
PT Isolated polynucleotide encoding a rice 5-enolpyruvylshikimate phosphate
PT synthase from rice is used for producing transgenic plants with
PT enhanced resistance to glyphosate herbicide
XX
PS Example 20; Page 55; 89pp; English.
XX
CC The present invention relates to an isolated polynucleotide encoding a
CC phosphate synthase (EPSPS) gene. Vectors containing the polynucleotide
CC used to produce plant tissue and fertile whole plants which are
CC substantially tolerant or resistant to glyphosate herbicide. The
CC herbicide and to produce a herbicidal target which is used in the
CC transgenic in vitro screening of potential herbicides
XX
SQ Sequence 1343 BP; 349 A; 344 C; 245 G; 405 T; 0 other;

Query Match 96.0%; Score 542; DB 21; Length 1343;
Best Local Similarity 98.4%; Pred. No. 9.2e-156;
Matches 549; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

Qy 1 GATCAAGTGCAAAGGTCCGCCTTGTTTCTCCTCTGTCTCTTGATCTGATTAATCTTTGTTT 60
Db 783 GATCAAGTGCAAAGGTCCGCCTTGTTTCTCCTCTGTCTCTTGATCTGATTAATCTTTGTTT 842

Qy 61 TATGATTCGTTGAGTAATTTTGGGGCAA----AGCTTCGTCCACACATTTTTTTTTCTCAATCA 116
Db 843 TATGATTCGTTGAGTAATTTTGGGGCAAAGCTAGCTTCGTCCACACATTTTTTTTTCTCAATCA 902

Qy 117 ACAGTGCCGCAGTGGCGCTGATCTTGTATGCTATCCTGCAATCTGGGTGAACTTATGTCT 176
Db 903 ACAGTGCCGCAGTGGCGCTGATCTTGTATGCTATCCTGCAATCTGCCTGAACTTATTTCT 962

Qy 177 TTTATATCCTTCACTACCATGAAAAGACTAGTAATCTTTCTCGATGTAACATCTCCAGC 236
Db 963 TTTATATCCTTCACTCCCATGAAAAGCCTAGTAATCTTTCTCGATGTAACATCGTCCAGC 1022

Qy 237 ACTGCTATTACCGTGTGGTCCATCCGACAGTCTGGCTGAACACATCATACGATATTGAGC 296
Db 1023 ACTGCTATTACCGTGTGGTCCATCCGACAGTCTGGCTGAACACATCATACGATATTGAGC 1082

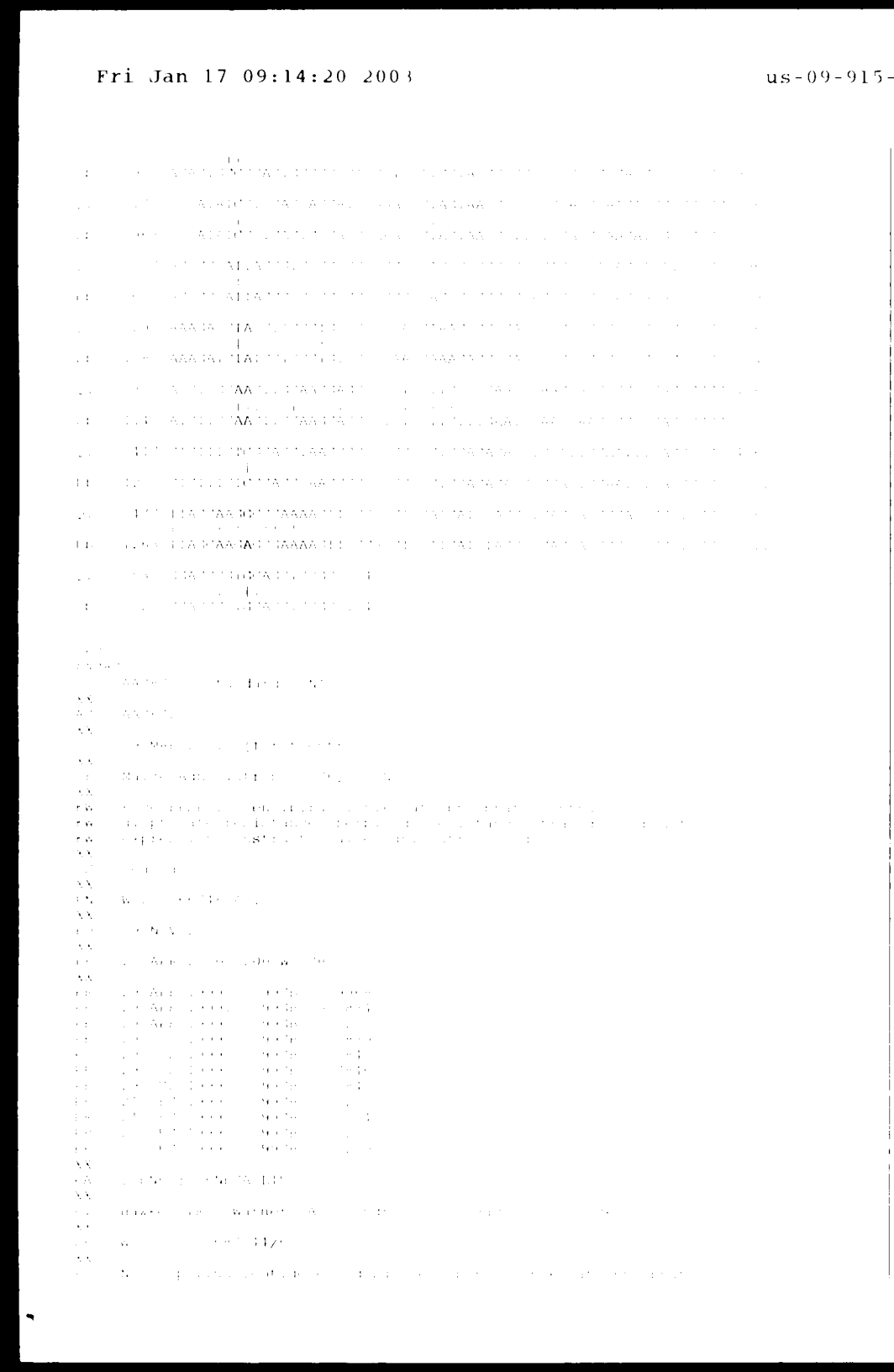
Qy 297 AAAGATCTATCTTCCCTGTTCTTTAATGAAAGACGTCATTTTCATCAGTATGATCTAAGA 356
Db 1083 AAAGATCTATCGTCGCTGTTCTTTAATGAAAGACGTCATTTTCATCAGTATGATCTAAGA 1142

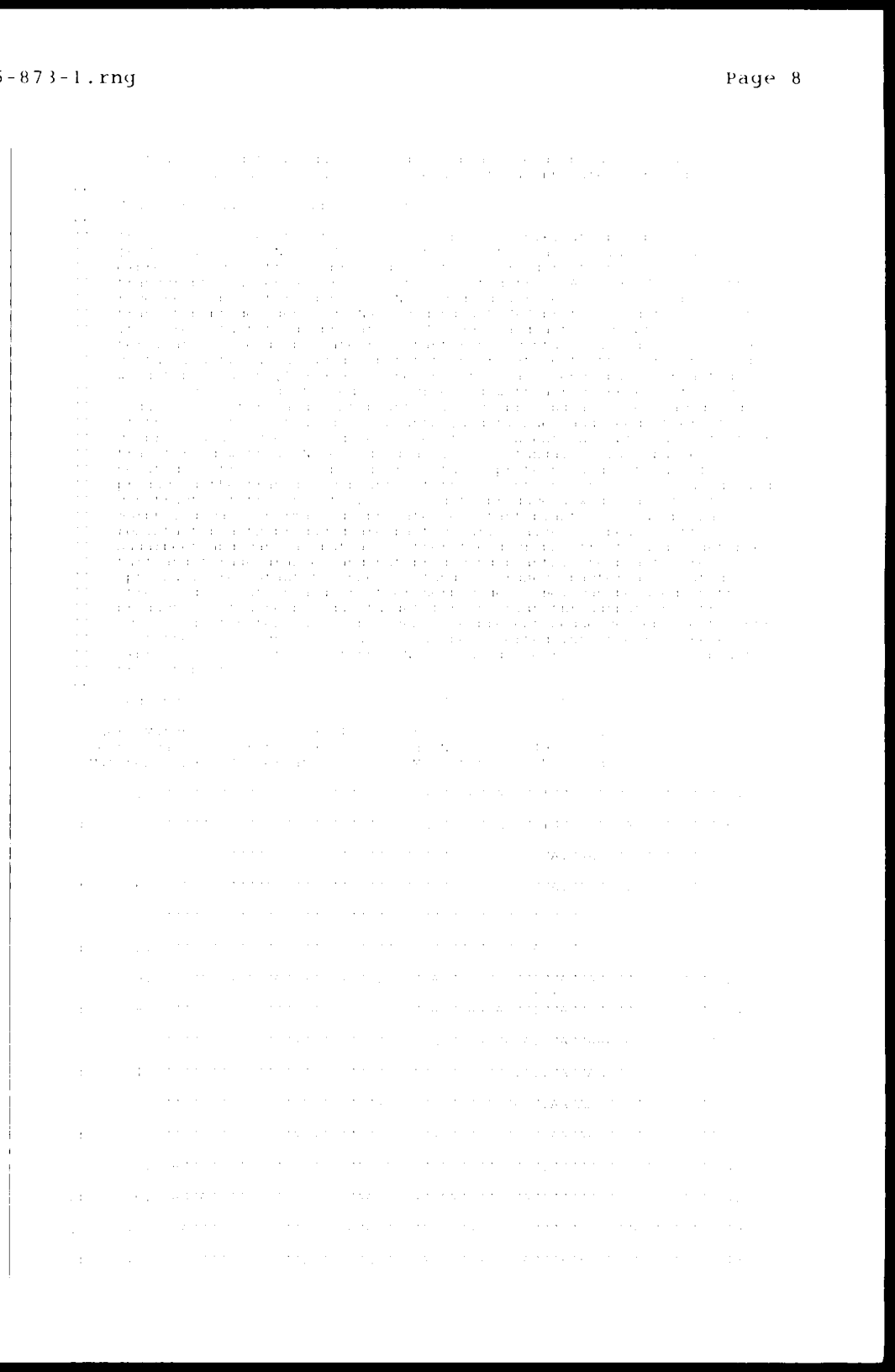
Qy 357 ATGTTGCAACTTGCAAGGAGGCGTTCTTCTTTGAATTTAACTAACTCGTTGACTCTT 416
Db 1143 ATGTTGCAACTTGCAAGGAGGCGTTCTTCTTTGAATTTAACTAACTCGTTGACTCTT 1202

Qy 417 CTGTTTCTCGGACGTAAGGCCTTTGCTGCTCCACACATGCCATGGAATTTTATCCTC 476
Db 1203 CTGTTTCTCGGACGTAAGGCCTTTGCTGCTCCACACATGCCATGGAATTTTATCCTC 1262

Qy 477 TTAGCAAGGGCGAAAAGTTTGCATCTTGATGATTAGCTCACTATCGATTGCTTCCT 536
Db 1263 TTAGCAAGGGCGAAAAGTTTGCATCTTGATGATTAGCTCACTATCGATTGCTTCCT 1322

Qy 537 GCATCCTCTGCCT 549
Db 1323 GCATCCTCTGCCT 1335





```
Qy    491 AAGTTTGCATCTTGATGATTTAGTTTGACTATCCGAT[illegible]
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 AAGTTTGCATCTTGATGATTTAGTTTGACTATCCGAT[illegible]
```

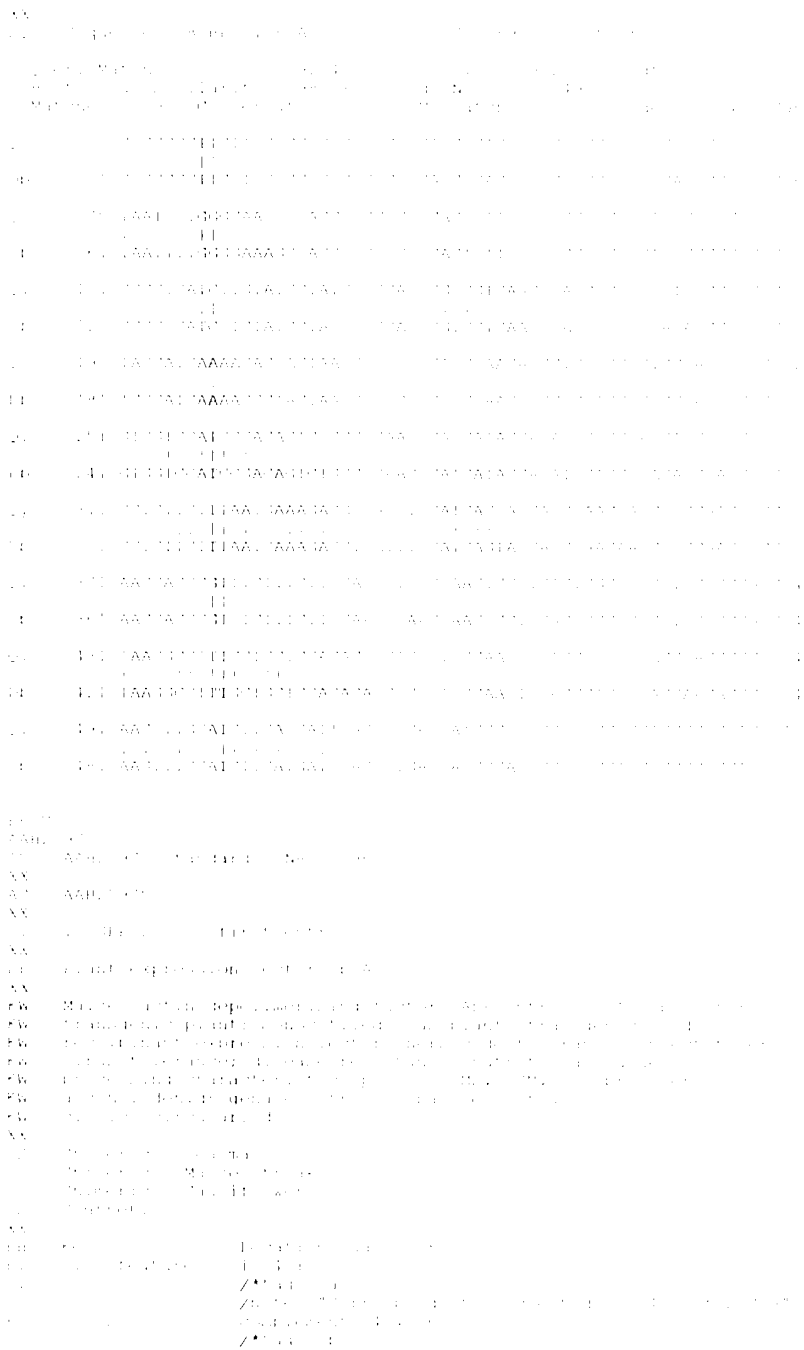
RESULT 11
AAC88405

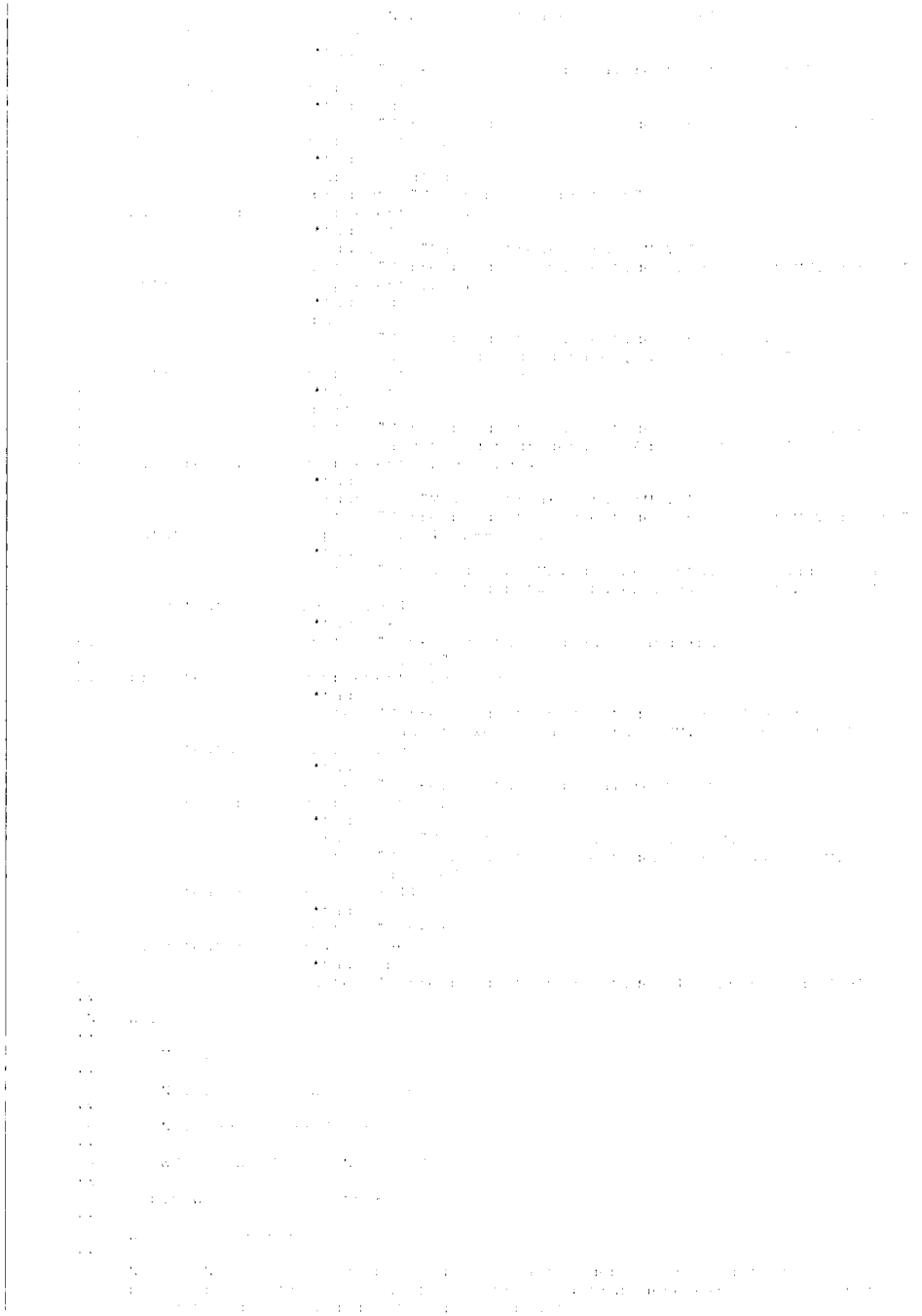
```
ID   AAC88405 standard; DNA; 538 BP.
XX
AC   AAC88405;
XX
DT   02-MAR-2001  (first entry)
XX
DE   Maize Adh1 intron.
XX
KW   Glyphosate; 5-enolpyruvylshikimate phosphate synthase; EPSPS;
KW   herbicide resistance; ss.
XX
OS   Zea mays.
XX
PN   WO200066746-A1.
XX
PD   09-NOV-2000.
XX
PF   20-APR-2000; 2000WO-GB01555.
XX
PR   29-APR-1999;   99GB-0009971.
PR   29-APR-1999;   99GB-0009972.
PR   29-JUL-1999;   99GB-0017837.
PR   29-JUL-1999;   99GB-0017842.
PR   21-DEC-1999;   99GB-0030190.
PR   21-DEC-1999;   99GB-0030206.
PR   21-DEC-1999;   99GB-0030214.
PR   21-DEC-1999;   99GB-0030216.
XX
PA   (ZENE ) ZENECA LTD.
XX
PI   Hawkes TR,  Warner SAJ,  Andrews CJ,  Bacham S,  Pickering AJ;
XX
DR   WPI; 2000-679763/66.
XX
PT   Novel polynucleotide encoding the rice 5-enolpyruvylshikimate phosphate
PT   synthase, used to produce glyphosate tolerant or resistant plants -
XX
PS   Claim 23; Page 56; 85pp; English.
XX
CC   The present invention relates to a glyphosate resistant rice
CC   5-enolpyruvylshikimate phosphate synthase (EPSPS) gene. This gene can
CC   be used to produce plant tissue and/or morphologically normal fertile
CC   whole plants which are tolerant or resistant to glyphosate herbicide,
CC   and in the production of a herbicidal target for the high throughput
CC   in vitro screening of potential herbicides.
XX
SQ   Sequence 538 BP; 118 A; 121 C; 110 G; 189 T; 0 other;

  Query Match             92.4%;  Score 512;  DB 21;  Length 538;
  Best Local Similarity   98.3%;  Pred. No. 1.1e-149;
  Matches  529;  Conservative    0;  Mismatches    [illegible];  Indels    [illegible];  Gaps    [illegible];

Qy     15 GTCCGCCTTGTTTCTCCTCGTGTCTCTTGATCTGACTAATCTTGGTTTATGATTCGTTGAG 74
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GTCCGCCTTGTTTCTCCTCGTGTCTCTTGATCTGACTAATCTTGGTTTATGATTCGTTGAG 60

Qy     75 TAATTTTGGGGAA----AGCTTCGTCCACAGTTTTTTTTTCGATGAACAGTGCCGCAGTG 13[illegible]
          |||||||||||||    |||||||||||||||||||||||||||||||||||||||||||
Db     61 TAATTTTGGGGAAAGCTAGCTTCGTCCACAGTTTTTTTTTCGATGAACAGTGCCGCAGTG 12[illegible]

Qy    131 GCGCTGATCTTGTATGCTATCCTGCAATCGTGGTGAACTTATTTCTTTTATATCCTTCACT 1[illegible]
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GCGCTGATCTTGTATGCTATCCTGCAATCGTGGTGAACTTATTTCTTTTATATCCTTCACT 1[illegible]

Qy    191 TACCATGAAAAGACTAGTAATCTTTCTCGATGTAACATCGTCCAGCACTGCTATTACCGT 25[illegible]
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    [illegible]
Qy    251 [illegible]
Db    241 [illegible]
Qy    311 [illegible]
Db    301 [illegible]
Qy    371 [illegible]
Db    361 [illegible]
Qy    431 [illegible]
Db    421 [illegible]
Qy    491 [illegible]
Db    481 [illegible]
```

RESULT 12
AAC89329

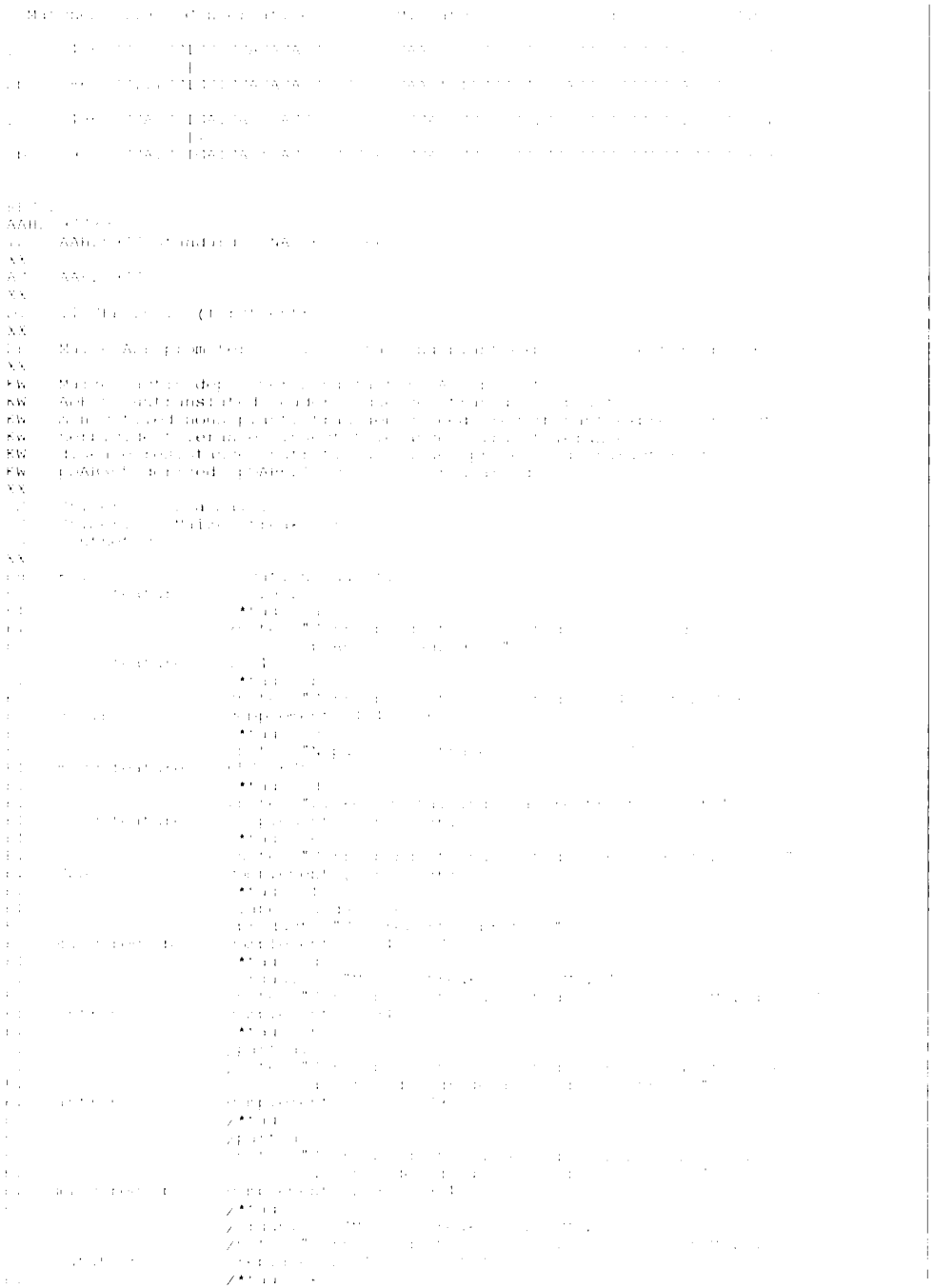
```
ID   AAC89329 standard; DNA; [illegible] BP.
XX
AC   AAC89329;
XX
DT   02-MAR-2001  (first entry)
XX
DE   Maize Adh1 intron.
XX
KW   [illegible] EPSPS;
KW   herbicide resistance; [illegible]
XX
OS   Zea mays.
XX
PN   WO2000[illegible]-A1.
XX
PD   09-NOV-2000.
XX
PF   20-APR-2000; 2000WO-[illegible]
XX
PR   29-APR-1999;   99GB-[illegible]
PR   29-APR-1999;   99GB-[illegible]
PR   29-APR-1999;   99GB-[illegible]
PR   29-APR-1999;   99GB-[illegible]
PR   29-APR-1999;   99GB-[illegible]
PR   29-JUL-1999;   99GB-[illegible]
PR   29-JUL-1999;   99GB-[illegible]
PR   21-DEC-1999;   99GB-[illegible]
PR   21-DEC-1999;   99GB-[illegible]
PR   21-DEC-1999;   99GB-[illegible]
XX
PA   (ZENE ) ZENECA LTD.
XX
PI   Hawkes TR, Warner SAJ, [illegible]
XX
DR   WPI; 2000-[illegible]/66.
XX
PT   Isolated polynucleotide [illegible]
PT   synthase from rice [illegible]
PT   enhanced resistance [illegible]
XX
PS   Claim 22; Page [illegible]; [illegible]pp; English.
XX
CC   The present invention relates to [illegible]
CC   phosphate synthase (EPSPS) gene. Vectors containing the gene [illegible]
CC   used to produce plant tissues and fertile whole plants which [illegible]
CC   substantially tolerant or substantially resistant to [illegible]
CC   herbicide and to produce a herbicidal target which [illegible]
CC   throughput in vitro screening of potential herbicides.
```





```
XX
PS   Example 1; Page 28-29; 42pp; English.
XX
CC   The invention relates to an isolated DNA molecule comprising a
CC   promoter or intron 1 of the maize actin depolymerising factor (ADF)
CC   gene (Zmbap). The ADF promoter constitutes bases 1-734 of the ADF gene
CC   fragment given in AAH25974 and intron 1 constitutes bases 882-2101 of
CC   AAH25974. The invention also relates to recombinant constructs comprising
CC   the maize ADF promoter or intron 1 operably linked to a heterologous
CC   nucleic acid sequence; transgenic plants, seeds or grain comprising
CC   such a construct; and methods of producing such plants or their seeds.
CC   The promoter and the intron of the maize ADF are useful in controlling
CC   transgene expression in both plants, particularly monocotyledonous
CC   plants, and preferably transgenic maize plants. Constructs of the
CC   invention can be used to express a variety of heterologous genes,
CC   particularly those which confer tolerance to herbicides, insects or
CC   viruses, and those that provide improved nutritional value or processing
CC   characteristics. The present sequence represents a plant expression
CC   vector, pDAB305, which was used as the basis for the construction of
CC   vectors (AAH25975 AAH25978) comprising the maize ADF promoter and 5'UTR
CC   and/or maize ADF intron 1.
XX
SQ   Sequence 5796 BP; 1405 A; 1489 C; 1411 G; 1491 T; 0 other;

  Query Match             21.5%;  Score 114;  DB 22;  Length 5796;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-26;
  Matches  114;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      4 GCATTTCTGCTCCACACATGTCCATTCGAATTTACGGTGTTTATAAAAAAGTT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2701 GCATTGCTGCTCCACACATGTCCATTCGAATTTACGGTGTTTATAAAAAAGTT 2645

Qy     496 TGCATCTTGATGATTTAGCTTGACTATGCGATTCCTTCCTCGATCGTTCAAAT 545
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2641 TGCATCTTGATGATTTAGCTTGACTATGCGATTGCTTCCTCGATCGTTCAAAT 2592


RESULT 14
AAV63741
ID   AAV63741 standard; DNA; 5897 BP.
XX
AC   AAV63741;
XX
DT   12-APR-1999  (first entry)
XX
DE   Plasmid pTGP190-1 containing 35S promoter and per5 5'UTR.
XX
KW   Peroxidase; per5 gene; maize; corn; transgenic plant; promoter;
KW   root; vector; plasmid pTGP190-1; CaMV; 35; cirsium; cynthia.
XX
OS   Chimeric - Escherichia coli.
OS   Chimeric - Zea mays.
OS   Chimeric - maize streak virus.
OS   Chimeric - Agrobacterium tumefaciens.
OS   Chimeric - Cauliflower mosaic virus.
OS   Chimeric - synthetic.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    18..20
FT                   /*tag= a
FT                   /note= "linker"
FT   promoter        31..282
FT                   /*tag= b
FT                   /note= "CaMV MCASTRAS nt 7093-7344"
FT   misc_feature    283..290
FT                   /*tag= c
FT                   /note= "linker"
FT   promoter        291..637
FT                   /*tag= d
FT                   /note= "CaMV MCASTRAS nt 7093-7439"
FT   misc_feature    638..657
FT                   /*tag= e
FT                   /note= "linker"
```

```
FT   intron          [illegible]
FT                   /*tag= f
FT                   /note= "[illegible]"
FT   misc_feature    [illegible]
FT                   /*tag= g
FT                   /note= "MCSV nt [illegible]"
FT   misc_feature    [illegible]
FT                   /*tag= h
FT                   /note= "MCSV nt [illegible]"
FT   intron          [illegible]
FT                   /*tag= i
FT                   /note= "ADF1.2 intron 1"
FT   misc_feature    [illegible]
FT                   /*tag= j
FT                   /note= "linker"
FT   misc_feature    [illegible]
FT                   /*tag= k
FT                   /note= "MCSV nt [illegible]"
FT   CDS             [illegible]
FT                   /*tag= l
FT                   /gene= "bar" [illegible]
FT   misc_feature    [illegible]
FT                   /*tag= m
FT                   /note= "polylinker [illegible]"
FT   CDS             [illegible]
FT                   /*tag= n
FT                   /note= "per5 coding sequence"
FT   misc_feature    [illegible]
FT                   /*tag= o
FT                   /note= "3' [illegible]"
XX
PN   WO9942594-A1
XX
PD   19-AUG-1999.
XX
PF   16-FEB-1999;  99WO-US003211.
XX
PR   17-FEB-1998;  98US-0074977.
XX
PA   (DEKA ) DEKALB GENETICS CORP.
XX
PI   Abbitt M;  Arnold GM;  Bo [illegible]
PI   Menke MA;  Petrolli [illegible]
XX
DR   WPI; 1999-488644/41.
XX
PT   New isolated nucleic acid sequence, useful for [illegible] plants [illegible]
PT   derived from the [illegible] per5 [illegible] gene
PT   (per5) gene
XX
PS   Example 20; Page 12; [illegible] pages; English.
XX
CC   This is the nucleotide sequence [illegible] a plant [illegible]
CC   cassette in which the [illegible] promoter and 5'UTR
CC   (which [illegible] the 35S [illegible] and the maize [illegible]
CC   proteinoid cationic peroxidase per5 gene (see AAV63740) 5'
CC   untranslated region (5'UTR). The 35S promoter [illegible]
CC   copies of the CaMV 35S [illegible] a maize [illegible]
CC   [illegible] and the [illegible] 5'UTR [illegible]
CC   per5 promoter [illegible] by the [illegible] 5'UTR, [illegible] which [illegible]
CC   by the [illegible] 5'UTR functions [illegible] the 35S [illegible]
CC   promoter. The invention [illegible] transformed cells,
CC   sequences, expression [illegible] and [illegible] per5
CC   maize per5 genes. [illegible] per5 [illegible]
CC   per5 regulatory sequences are useful [illegible] plants, [illegible]
CC   recombinant [illegible] transformed plants, per5 [illegible]
XX
SQ   Sequence 5897 BP; [illegible]

  Query Match             [illegible];  Score [illegible];  DB [illegible];  Length [illegible];
  Best Local Similarity   [illegible];  Pred. No. [illegible];
```



Search completed: January 17, 2003, 05:22:17
Job time : 307 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: January 17, 2003, 05:11:47 ; Search time 2224 Seconds
(without alignments)
4034.207 Million cell updates/sec

Title: US-09-915-873-1
Perfect score: 554
Sequence: 1 gatcaaatqcaaaaqtcaqc.....ctaqaac qtqcaqctacaq 554

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8697743476 residues

Total number of hits satisfying chosen parameters: 42308152

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: qb_est1:*
10: qb_est2:*
11: qb_htc:*
12: qb_est3:*
13: qb_est4:*
14: qb_est5:*
15: em_estfun:*
16: em_estom:*
17: qb_qss:*
18: em_qss_hum:*
19: em_qss_inv:*
20: em_qss_pln:*
21: em_qss_vrt:*
22: em_qss_fun:*
23: em_qss_mam:*
24: em_qss_mus:*
25: em_qss_other:*
26: em_qss_pro:*
27: em_qss_rod:*

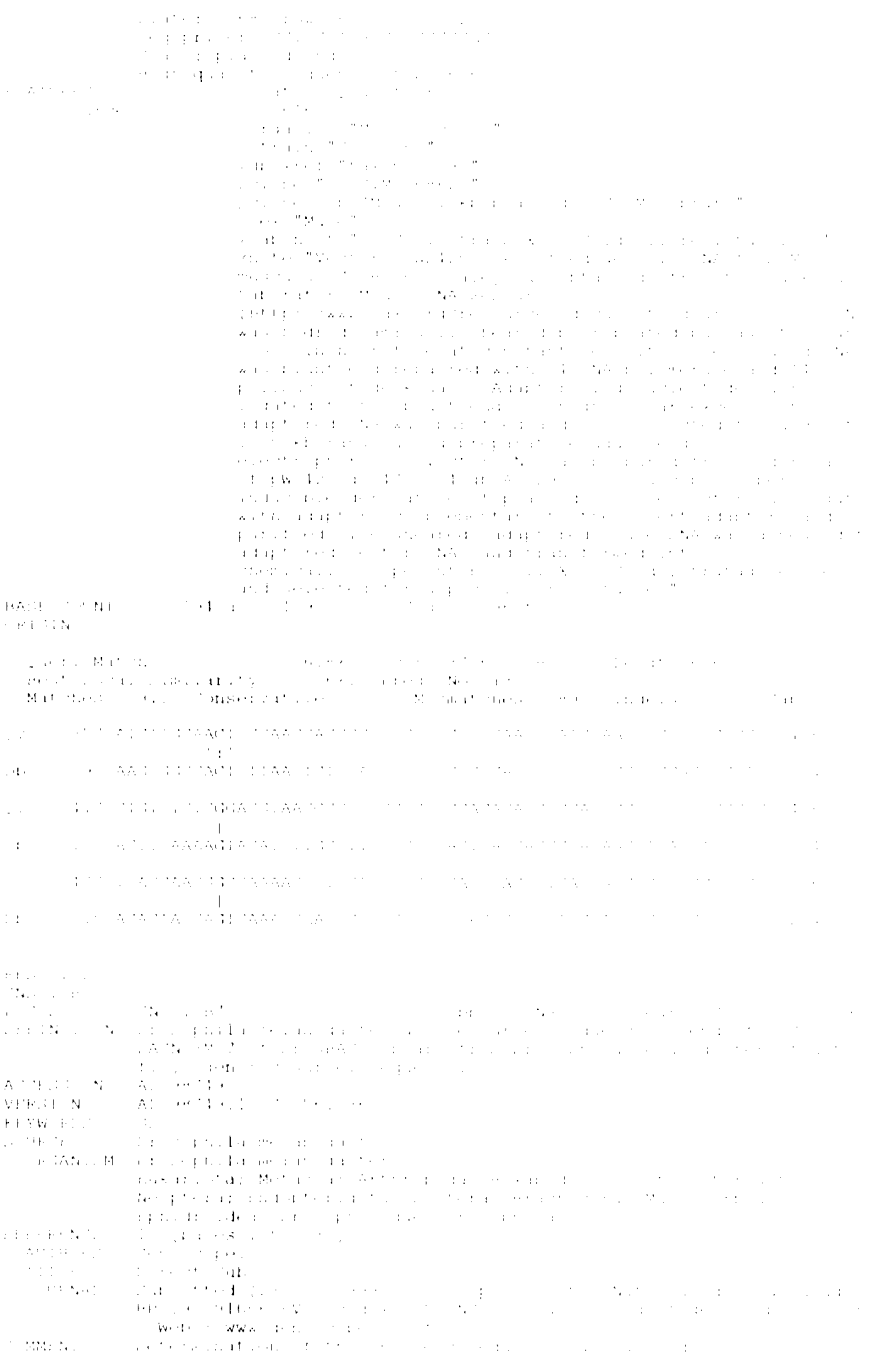
Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

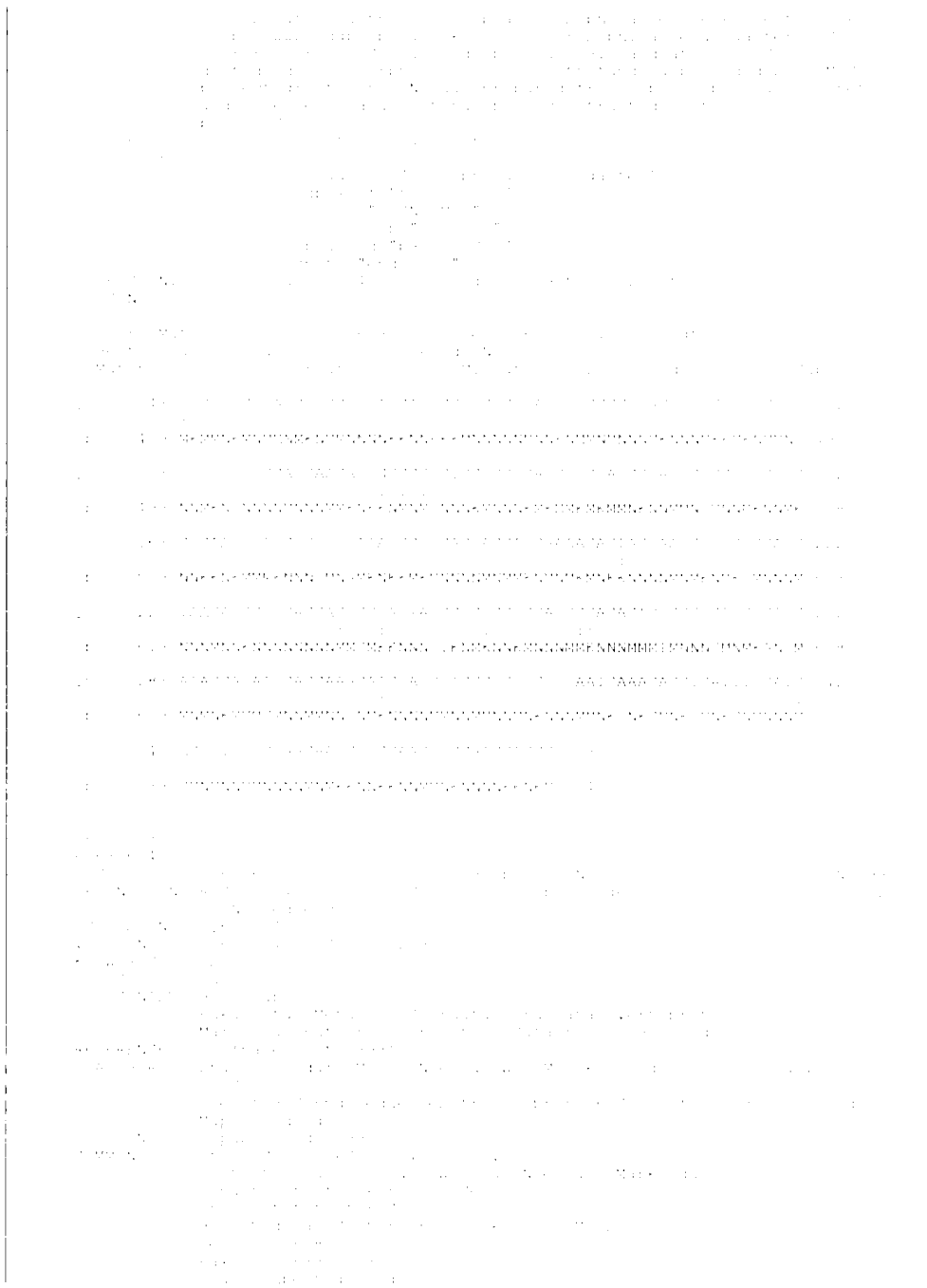
SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	37.6	6.8	616	17	AZ337633	AZ337633 EM:6
2	37.6	6.8	787	17	CNS010B7	AL098749 Drosophil
3	37	6.7	398	17	AQ645634	AQ645634 RPCI-11-4
c 4	37	6.7	402	9	AA773011	AA773011 ucd4pclu_s
5	37	6.7	546	17	AZ623239	AZ623239 EM:6
c 6	36.8	6.6	1101	17	CNS00E2D	AL067697 Drosophil
c 7	36.8	6.6	1101	17	CNS00E2D	AL067697 Drosophil
8	36.4	6.6	181	12	BE731[illegible]	BE731[illegible] BE
c 9	36.2	6.5	891	11	B[illegible]4	B[illegible]
c 10	36.2	6.5	917	11	B[illegible]2	B[illegible]
c 11	36	6.5	516	17	A[illegible]	A[illegible]
12	36	6.5	540	17	A[illegible]	A[illegible]
c 13	36	6.5	613	17	B[illegible]	B[illegible]
c 14	36	6.5	1101	17	CNS[illegible]	A[illegible]
c 15	35.8	6.5	699	17	B[illegible]	B[illegible]
c 16	35.6	6.4	99	17	CNS[illegible]	A[illegible]
c 17	35.4	6.4	695	9	BE[illegible]	BE[illegible]
c 18	35.2	6.4	999	9	BE[illegible]	BE[illegible]
19	35.2	6.4	599	14	A[illegible]	A[illegible]
20	35.2	6.4	601	17	AZ[illegible]	AZ[illegible]
c 21	35	6.3	595	17	AZ[illegible]	AZ[illegible]
22	35	6.3	599	14	B[illegible]	B[illegible]
23	35	6.3	992	17	CNS[illegible]	A[illegible]
24	35	6.3	1101	17	CNS[illegible]	A[illegible]
25	35	6.3	1102	17	CNS[illegible]	A[illegible]
26	35	6.3	1492	17	AL[illegible]	AL[illegible]
c 27	34.8	6.3	150	17	A[illegible]	A[illegible]
28	34.8	6.3	599	17	A[illegible]	A[illegible]
29	34.8	6.3	714	17	A[illegible]	A[illegible]
c 30	34.6	6.2	128	17	A[illegible]	A[illegible]
31	34.6	6.2	599	17	A[illegible]	A[illegible]
c 32	34.6	6.2	621	17	AC[illegible]	AC[illegible]
33	34.6	6.2	704	11	B[illegible]	B[illegible]
c 34	34.4	6.2	862	17	CNS[illegible]	A[illegible]
c 35	34.4	6.2	1101	17	CNS[illegible]	A[illegible]
c 36	34.4	6.2	1201	17	CNS[illegible]	A[illegible]
37	34.4	6.2	599	14	BM[illegible]	BM[illegible]
38	34.4	6.2	601	17	A[illegible]	A[illegible]
39	34.4	6.2	695	14	BM[illegible]	BM[illegible]
40	34.4	6.2	699	14	BM[illegible]	BM[illegible]
c 41	34.4	6.2	1201	17	CNS[illegible]	A[illegible]
c 42	34.2	6.2	599	17	BE[illegible]	BE[illegible]
c 43	34.2	6.2	601	17	CNS[illegible]	AZ[illegible]
44	34.2	6.2	1101	17	AC[illegible]	AC[illegible]
45	34	6.1	201	4	AA[illegible]	AA[illegible]

ALIGNMENTS

RESULT 1
AZ337633
LOCUS AZ337633 [illegible] DNA [illegible]
DEFINITION [illegible]
ACCESSION AZ337633
VERSION AZ337633.1 [illegible]
KEYWORDS "BSS."
SOURCE [illegible]
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; [illegible]
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 616)
AUTHORS Dumas,L., Aoyagi,A., Barber,M., Bonfield,J., [illegible]
Ishaq,H., Johnstone,L., Matthews,L., Mott,R., Rogers,J., [illegible]
Mu,F., Rosen,B., Rosenthal,A., Shownkeen,R., Ustaszewska,A. and [illegible]
and Williams,L.
TITLE Mouse whole genome shotgun [illegible]
Possible inserts
JOURNAL Unpublished (2002)
COMMENT Contact: Robert W. Weiss
[illegible]
[illegible]
Rm. 409, Bioengineering Program, Research Plaza, [illegible]
84112, USA
Tel: 801 585 5000
Fax: 801 585 7167
Email: [illegible]
Insert length: [illegible]

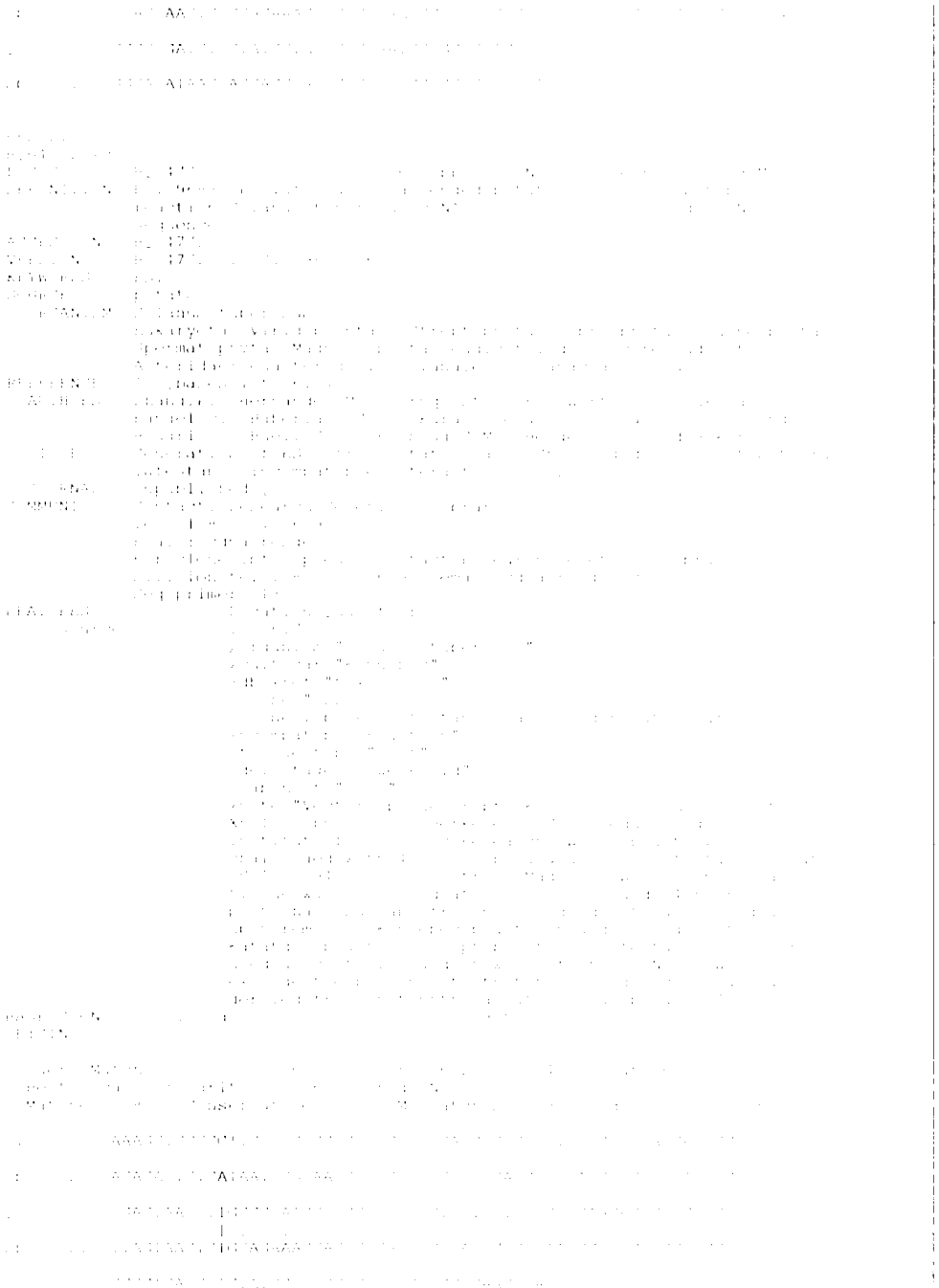


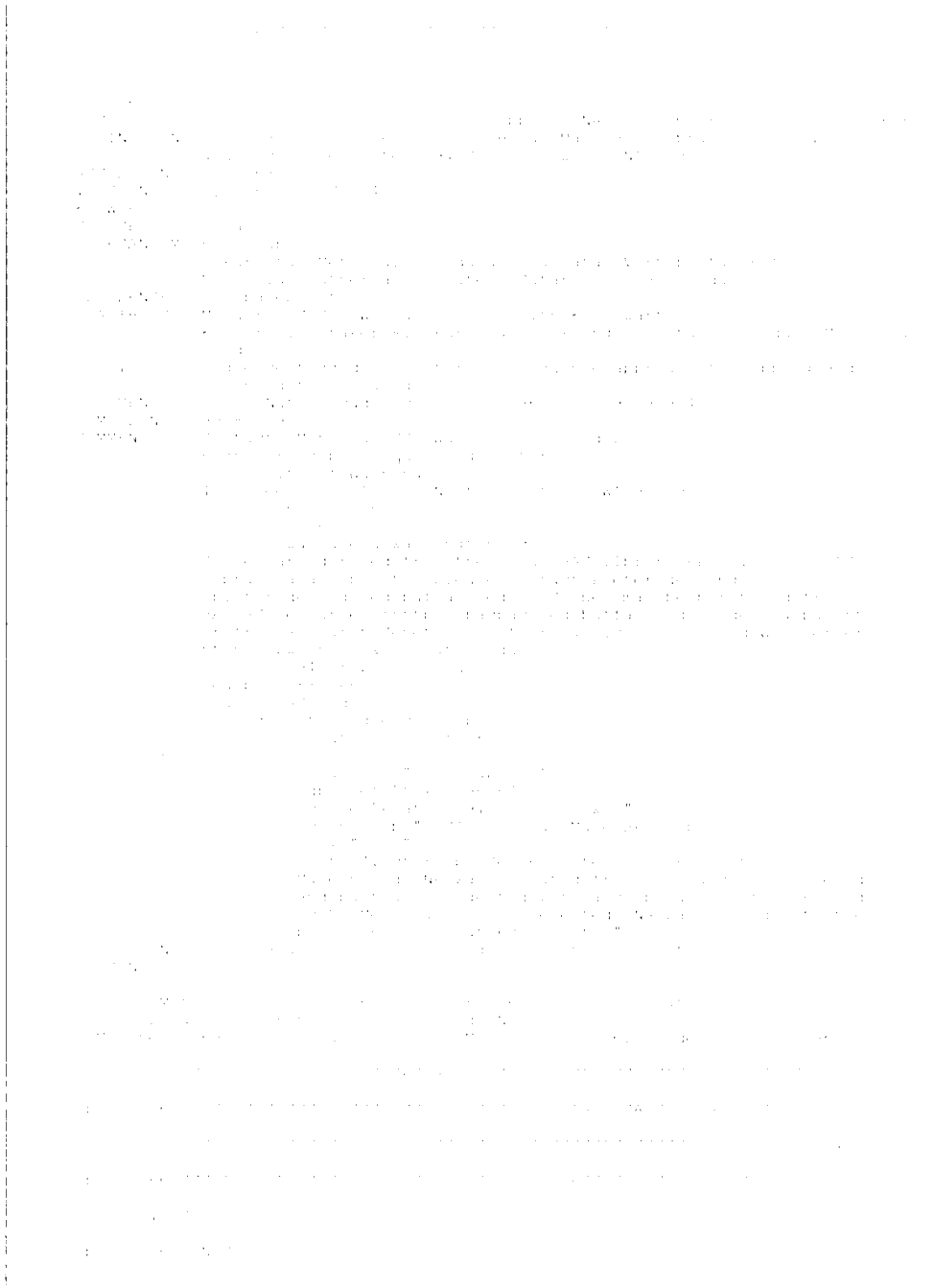


```
            Clones are derived from the human BAC library RPCI 11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/) or from
            Research Genetics (info@resgen.com). BAC end search page:
            http://www.tigr.org/cgi-bin/BlastSearch/blast_bac_end_seq.cgi
            Seq primer: SP6
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1..398
                     /organism="Homo sapiens"
                     /db_xref="GDB:7684124"
                     /db_xref="taxon:9606"
                     /clone="RPCI-11-480H21"
                     /clone_lib "RPCI-11"
                     /sex="Male"
                     /cell_type="Lymphocytes"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     RPCI11 Human Male BAC Library"
BASE COUNT       92 a    128 c     74 g    104 t
ORIGIN

  Query Match             6.7%;  Score 37;  DB 12;  Length 398;
  Best Local Similarity   56.6%;  Pred. No. 15;
  Matches   70;  Conservative    0;  Mismatches   55;  Indels    9;  Gaps    9;

Qy     375 AGGCGTTTCTTTCTTTGAATTTAAATAACTCTTTGACTCTCCTTTTTCTTTAATAAAC 434
           ||||  ||| | |    || |  ||   || || | | |||| ||| |  | |  |||
Db     190 AGGCCACTCTGTTTGAACAIGTCTCTGTGTCTTTAAATAGCCTTCTTTCCCACTGAAG 249

Qy     435 GCCTTTGCTGCTCCACACAIGTCCATTCGAATTTTACCGTGTTTAGCAAGGGCGAAAAGC 494
           ||| | |  | |  |||   | |  |   ||  |||  | | |||  |   | | | |||
Db     250 GCCTGTCCCTCCCACCACGAATAAACATGATCTTTGTCTTCTTAAAAC-TCCCATAAGC 308

Qy     495 TTGCA 499
           |||||
Db     309 TTGCA 313


RESULT 4
AA773011/c
LOCUS       AA773011                 402 bp    mRNA    linear   EST 29-JAN-1998
DEFINITION  ac48c10.s1 Stratagene LNT cDNA Homo sapiens cDNA clone
            IMAGE:859698 3', mRNA sequence.
ACCESSION   AA773011
VERSION     AA773011.1  GI:2824582
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 402)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,J., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
            ,J., Moore,B., Scheinberg,K., Stoptoe,M., Tan,F., Theising,P.,
            White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 387.
FEATURES             Location/Qualifiers
     source          1..402
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:859698"
                     /clone_lib "LLNL cDNA Library"
                     /sex="Male"
                     /cell_type="Lymphocyte"
                     /note="Vector: pSPORT1; Site_1: NotI; Site_2: SalI;
                     Stratagene LNT cDNA Library (#937233)"
BASE COUNT       91 a    102 c     94 g    115 t
ORIGIN

  Query Match             6.7%;  Score 37;  DB 12;  Length 402;
  Best Local Similarity   56.6%;  Pred. No. 15;
  Matches   70;  Conservative    0;  Mismatches   55;  Indels    9;  Gaps    9;

Qy     375 AGGCGTTTCTTTCTTTGAATTTAAATAACTCTTTGACTCTCCTTTTTCTTTAATAAAC 434
           ||||  ||| | |    || |  ||   || || | | |||| ||| |  | |  |||
Db     190 AGGCCACTCTGTTTGAACAIGTCTCTGTGTCTTTAAATAGCCTTCTTTCCCACTGAAG 249

Qy     435 GCCTTTGCTGCTCCACACAIGTCCATTCGAATTTTACCGTGTTTAGCAAGGGCGAAAAGC 494
           ||| | |  | |  |||   | |  |   ||  |||  | | |||  |   | | | |||
Db     250 GCCTGTCCCTCCCACCACGAATAAACATGATCTTTGTCTTCTTAAAAC-TCCCATAAGC 308

Qy     495 TTGCA 499
           |||||
Db     309 TTGCA 313


RESULT 5
AZ633xxx
LOCUS       AZ633xxx                 xxx bp    DNA     linear   EST 29-JAN-2001
DEFINITION  IMAGE:xxxxxx xxx, mRNA sequence.
ACCESSION   AZ633xxx
VERSION     AZ633xxx.1  GI:xxxxxxx
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 516)
  AUTHORS   Marra,M., Asano,J.A., Barber,M., Bonaldo,M.F., Brent,M., Brown,S.,
            Castle,A., Chinwalla,A., Delehaunty,K., Dewar,K., Dietrich,N.,
            Fulton,L., Gibbs,R., Hillier,L., Jackson,S., Kucaba,T., Lennon,G.,
            Mardis,E., Martin,J., Minx,P., Morgan,M., Rohlfing,T., Rudd,K.,
            Schein,J., Schmidt,H., Sekhon,M., Shah,N., Smith,S., Spieth,J.,
            Strong,S., Tan,F., Tyson,R., Waterston,R., Wohldmann,P., Wylie,T.
            and Wilson,R.
  TITLE     Mouse unfinished genomic sequence from chromosome 11, clone
            RP23-xxxx
  JOURNAL   Unpublished (xxxx)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: SP6
            Class: BAC ends.
            High quality sequence start: 1
            High quality sequence stop: xxx.
FEATURES             Location/Qualifiers
     source          1..xxx
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:xxxx"
                     /clone="RP23",Mouse xxx"
                     /clone_lib "Mouse xxx BAC Library"
                     /sex="Male"
                     /tissue_type="xxx xxx xxx"
                     /note="Vector: xxx; Site_1: xxx; Site_2: xxx;
                     Mouse xxx (male) xxx Library"
```

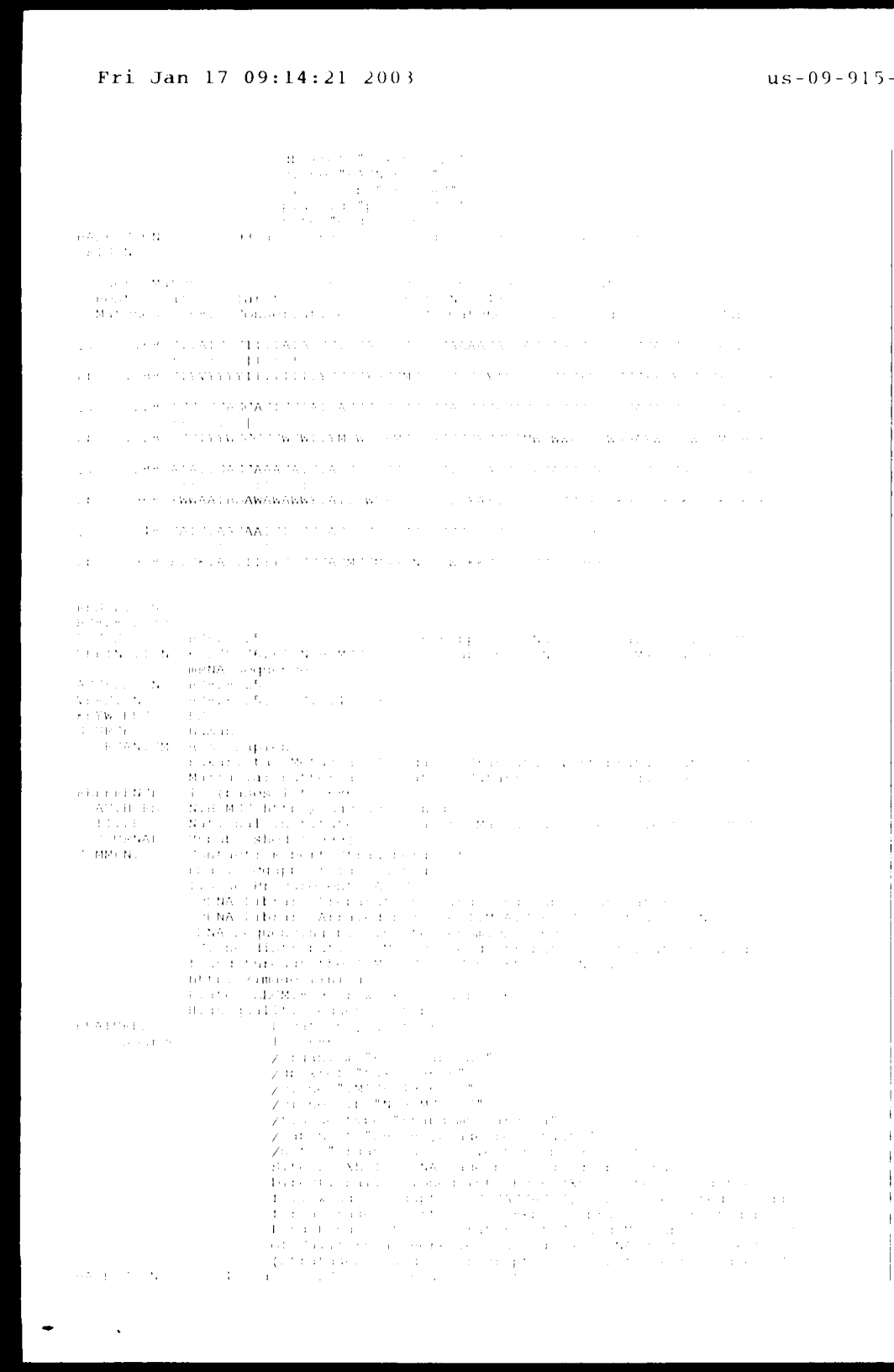
```
ORIGIN

  Query Match              6.6%;  Score 36.8;  DB 12;  Length 21 7;
  Best Local Similarity   51.9%;  Pred. No. 39;
  Matches   83;  Conservative    0;  Mismatches   77;  Indels    0;  Gaps    0;

Qy      33 TCTTCTCTTCATCTGACTAATCTTCGTTCATCATCTCATTCAATCTTCTTCAATTCTT 91
           | ||||| ||  | |       | ||| ||| |  || |    ||  || || |       |
Db     979 TGTGTCTGTTTCTGTCTTCTTTTTTGTTTTTTCCTTTGCGCCGTGGTCTCTCTGTTTCTT 920

Qy      92 TCGTCCACAGTTTTTTTTTCAATGAACAGTCCCTTACTGCGGTTCAATCTCGTAATCTAAC 151
           ||   |||  |||| |  |  ||    ||||     ||    ||  | | | | | | | |
Db     919 TCTCTCACTCTTTGTCGTGCTTGCGTTGTGCTTGCGTTTGTCGTTTTTTTTTTTTTTT 86

Qy     152 CTGCAATCGTGGTGAACTTATGTCTTTTATATCGTTCAAT 190
           ||    ||||| ||  ||  |||||| | | ||  ||| |
Db     859 CTTGTCTCGTGTTGCTCTGTTGTCTTGTTTGTCTGTCAAT 820


RESULT 8
BF744363
LOCUS       BF744363                 461 bp    mRNA    linear   EST 16-JAN-2001
DEFINITION  CM3-BT0861-281000-429-b02 BT0861 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF744363
VERSION     BF744363.1  GI:12071859
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 461)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.R.,
            Brunstein,A., de Oliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the PARSE/LICR Human Cancer Genome
            Project.  This entry can be seen in the following URL
            (http://www.ludwig.edu.br/cgi-bin/db/dbRun?pathway=no&project=
            281000-429-b02&t3=2000 10-28&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 17
            High quality sequence stop: 458.
FEATURES             Location/Qualifiers
     source          1. .461
                     /organism="Homo sapiens"
                     /db_xref "taxon:9606"
                     /clone_lib "BT0861"
                     /dev_stage "Adult"
                     /note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2:
                     SmaI; A mini-library was made by cloning products boiled
                     from  ORESTES PCR (U.S. Letters Patent application No. 10
                     ,716 -  Ludwig Institute for Cancer Research) profiles
                     into  the pUC 18 vector. Reverse transcription of tissue
                     mRNA and cDNA amplification were performed under low
                     stringency conditions."
BASE COUNT      105 a     76 c     69 g    211 t
ORIGIN

  Query Match              6.6%;  Score 36.8;  DB 12;  Length 461;
```

```
  Best Local Similarity   [illegible]
  Matches   [illegible]

Qy      32 [illegible]
           [illegible]
Db     [illegible]

Qy      92 [illegible]
           [illegible]
Db      99 [illegible]

Qy     152 [illegible]
           [illegible]
Db     [illegible]


RESULT 9
[illegible]
LOCUS       [illegible]              [illegible] bp    mRNA    [illegible]
DEFINITION  [illegible]
            [illegible]
            [illegible], mRNA sequence.
ACCESSION   [illegible]
VERSION     [illegible]
KEYWORDS    EST.
SOURCE      [illegible]
  ORGANISM  [illegible]
            [illegible]
            [illegible]
            [illegible]
REFERENCE   1  (bases 1 to [illegible])
  AUTHORS   [illegible]
            [illegible]
            [illegible]
  TITLE     [illegible]
  JOURNAL   Unpublished (2000)
COMMENT     Other_seqids: [illegible]
            Contact: Wilson Davis
            The Institute for Genomic Research
            9712a Medical Center Dr, Rockville, MD 20850, USA
            Email: gotc@tigr.org
            This clone is available [illegible] for the Research Community
            Research Genetics [illegible]
            cdna-request@
            Seq primer: T7.
FEATURES             Location/Qualifiers
     source          1. .554
                     /organism="Solanum tuberosum"
                     /db_xref="taxon:4113"
                     /clone="[illegible]"
                     /clone_lib="[illegible]
                     [illegible]
                     /tissue_type="mixed tissues"
                     /lab_host="[illegible]"
                     /note="Vector: pBluescript SK( ); Site_1: [illegible]
                     Xhoi; Supplier: [illegible]
                     Intestinal Stromal Tumors [illegible]
                     Library Type: [illegible] normalized [illegible]
                     [illegible]"
BASE COUNT      [illegible]
ORIGIN

  Query Match              [illegible]
  Best Local Similarity   [illegible]
  Matches   [illegible]

Qy      [illegible]
           [illegible]
Db     [illegible]

Qy      [illegible]
           [illegible]
```





```
RESULT 12
AQ483179
LOCUS       AQ483179                 586 bp    DNA     linear   GSS 24-APR-1997
DEFINITION  RPCI-11-241P4.TV RPCI-11 Homo sapiens genomic clone RPCI-11-241P4,
            DNA sequence.
ACCESSION   AQ483179
VERSION     AQ483179.1  GI:4670583
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 586)
  AUTHORS   Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and Venter
            ,J.C.
  TITLE     Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genetics (info@resgen.com). BAC end search page:
            http://www.tigr.org/db/humgen/bac_end/bac_end_search.html.
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1..586
                     /organism="Homo sapiens"
                     /db_xref="GDB:7592523"
                     /db_xref="taxon:9606"
                     /clone="RPCI-11-241P4"
                     /clone_lib "RPCI-11"
                     /sex="Male"
                     /cell_type "Lymphocytes"
                     /note="Vector: pBACe3.6; site_1: EcoRI; site_2: EcoRI;
                     RPCI11 Human Male BAC Library"
BASE COUNT      178 a    122 c    104 g    180 t      2 others
ORIGIN

  Query Match             6.5%;  Score 36;  DB 12;  Length 586;
  Best Local Similarity   60.6%;  Pred. No. 35;
  Matches   57;  Conservative    0;  Mismatches   37;  Indels    0;  Gaps    0;

Qy     128 GTGGCGCTGATCTTGTATGCTATCCTGCAATCGTGGTGAACTTATGTCTTTTATATATTTT 187
           || ||  |  |  ||||||  || ||||||||| | |  || | | | | |  |  | ||
Db     464 GTAGCTTTTCTTCTGTATGAGATTCTGCAATCGTAGACTCTTTCTNTATATTATTTCATT 523

Qy     188 CACTACCATGAAAAGACTAGTAATCTTTCTCGAT 221
           ||| |   | ||||  |   | | | | || ||
Db     524 CACAAATTTCAAAATTCACANAGTCTCTATCAAT 557


RESULT 13
BG963011/c
LOCUS       BG963011                 847 bp    mRNA    linear   EST 13-JUN-2001
DEFINITION  602828917T1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4982009 5',
            mRNA sequence.
ACCESSION   BG963011
VERSION     BG963011.1  GI:14350648
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 847)
  AUTHORS   NIH MGC (http://mgc.nci.nih.gov)
  TITLE     National Institutes of Health Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg
            Email: cgapbg@mail.nih.gov
            Tissue Procurement: Robert L. Strausberg, M.D.
            cDNA Library Preparation: Soares et al.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Baylor College of Medicine
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Note: CLAM=089  [exact]  [vector]
            High quality sequence stop: 847.
FEATURES             Location/Qualifiers
     source          1..847
                     /organism="Mus musculus"
                     /strain="BY/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4982009"
                     /clone_lib "NCI_CGAP_Co24"
                     /tissue_type "Mammary gland"
                     /note="Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI;
                     Seq primer: SP6 universal; Site_1 Bases: 1-8 Kb; Constructed
                     Sequence: NotI-SalI; Primer: a NCI_CGAP_Co24 primer"
BASE COUNT      [illegible]
ORIGIN

  Query Match             [illegible]  Score 36;  DB 12;  Length 847;
  Best Local Similarity   64.3%;  Pred. No. 1.2;
  Matches   54;  Conservative    0;  Mismatches   30;  Indels    0;  Gaps    0;

Qy      [illegible]
           ||| ||| ||  |||| ||  | | |  | |||  |  | |  |  |  |  |  |  | |
Db    536 CTGCTCTTTTTTTTTTGTGCTTTTCTTCATTTTGTTTTTTTTTTTTTTTTTTTTTTTTTTT 595

Qy      [illegible]
           ||| |      || |||| ||
Db    476 TTCTTTTTTTTTTTTTTTTTTTTTTTTTTT 506


RESULT 14
CNS01GSB/c
LOCUS       CNS01GSB                 [illegible] bp    DNA     [illegible]
DEFINITION  Drosophila melanogaster BAC RP98-[illegible]A12 complete sequence.
            BACR[illegible]A12 of the BACR [illegible] Drosophila melanogaster
            [illegible] genomic survey sequence.
ACCESSION   AL[illegible]
VERSION     AL[illegible].1  [illegible]
KEYWORDS    HTG.
SOURCE      Drosophila melanogaster
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to [illegible])
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted ([illegible]) Genoscope - Centre National de Sequencage,
            BP 191 91006 EVRY cedex - FRANCE (E-mail: seq@[illegible].fr, Tel:
            - WEB: www.genoscope.cns.fr)
COMMENT     Determination of this BAC clone sequence was carried out in
            collaboration with the European Drosophila Genome Project (EDGP),
            http://www.edgp.imcb.nus.edu.sg. This Drosophila melanogaster BAC
            library (BACR BAC) was made by Alain Billault at CNRS (Centre
            d'Etude du Polymorphisme Humain) with funding from the EC
            project grant. The DNA was prepared from male flies, Canton S
            and Genetic Force 2. The host strain is [illegible]
            phoDBA2[illegible].
FEATURES             Location/Qualifiers
     source          1..[illegible]
                     /organism="Drosophila melanogaster"
```



GenCore version 5.1.3

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 17, 2003, 06:58:42 ; Search time 4.7 Seconds
(without alignments)
946.827 Million cell updates/sec

Title: US-09-915-873-3
Perfect score: 178
Sequence: 1 acccatgcaaatgctgaaaca.....tcaaaaaaaaaaaaaaaaa 96

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 225784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV-xlh
-Q=/cgn2_1/tmp/pto sp_9[4]_US-09-915-873_3.fasta_[4012003_171632_17282] -NOPG -LOPOP
-DB=SwissProt_40 -QFMT=fastan -SUFFIX=n2p.rsp -MINMATCH=0.1 -LOOPCL 0
-LOOPEXT=0 -UNITS=bits -START 1 -END= 1 MATRIX BLOSUM62 TRANS=human4v.cdi
-LIST=45 -DOCALIGN 200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN 15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -REAPSIZE=5e8 -MINLEN=0 -MAXLEN 2000000000
-USER=US09915873_@CGN_1_1_4_#runat_14012003_171632_17282 -NOPG -LOPO
-NO_XLPXY -NO_MMAP -IARGP@UERY -NET_SCORES 0 -WAIT -DOCSUM -ENV_TIME 01:12
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP 10 -XGAEXT=0.5 -YGAPOP 10 -YGAEXT 7
-YGAPOP=10 -YGAEXT=0.5 -DELOP=6 -DELEXT 7

Database : SwissProt_40:*

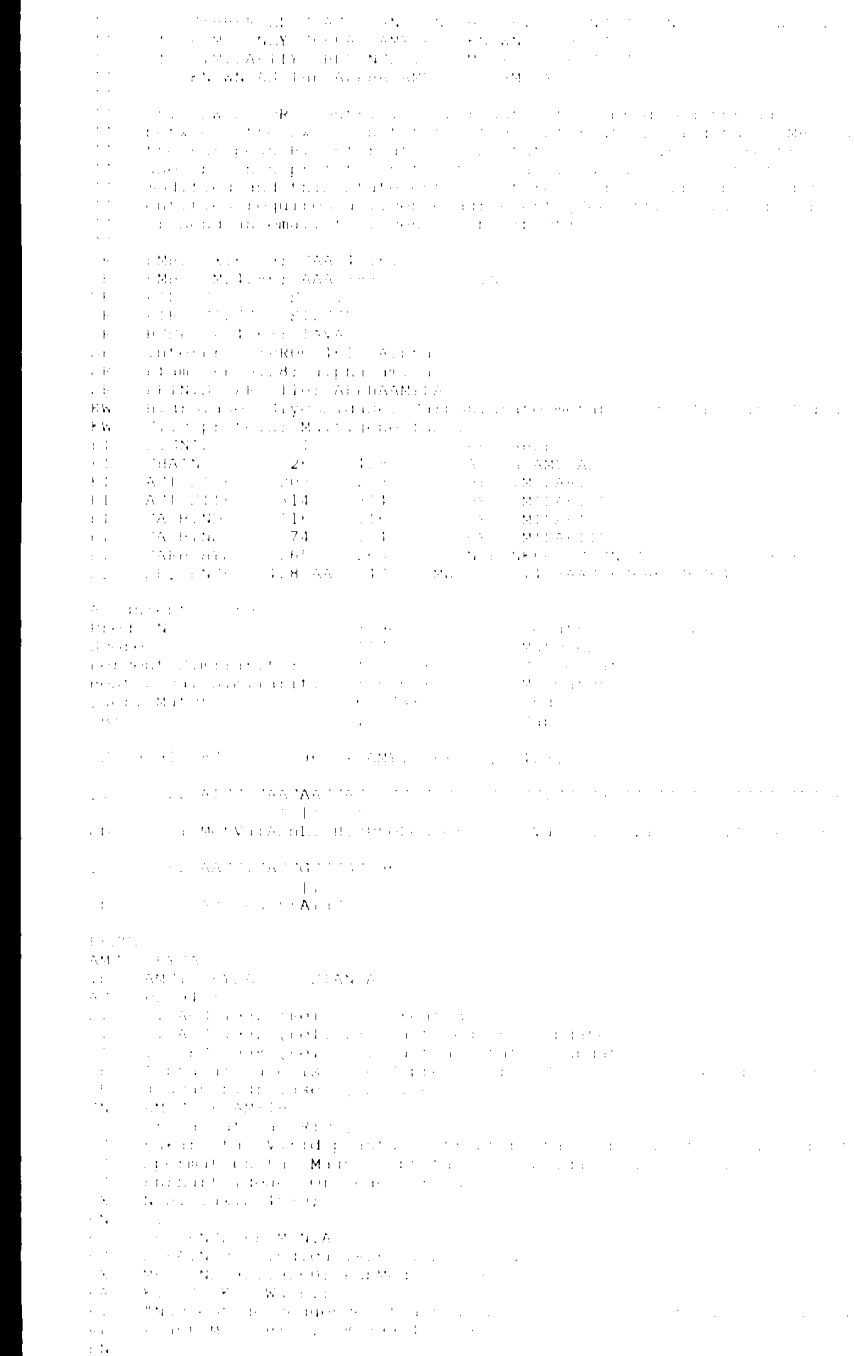
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

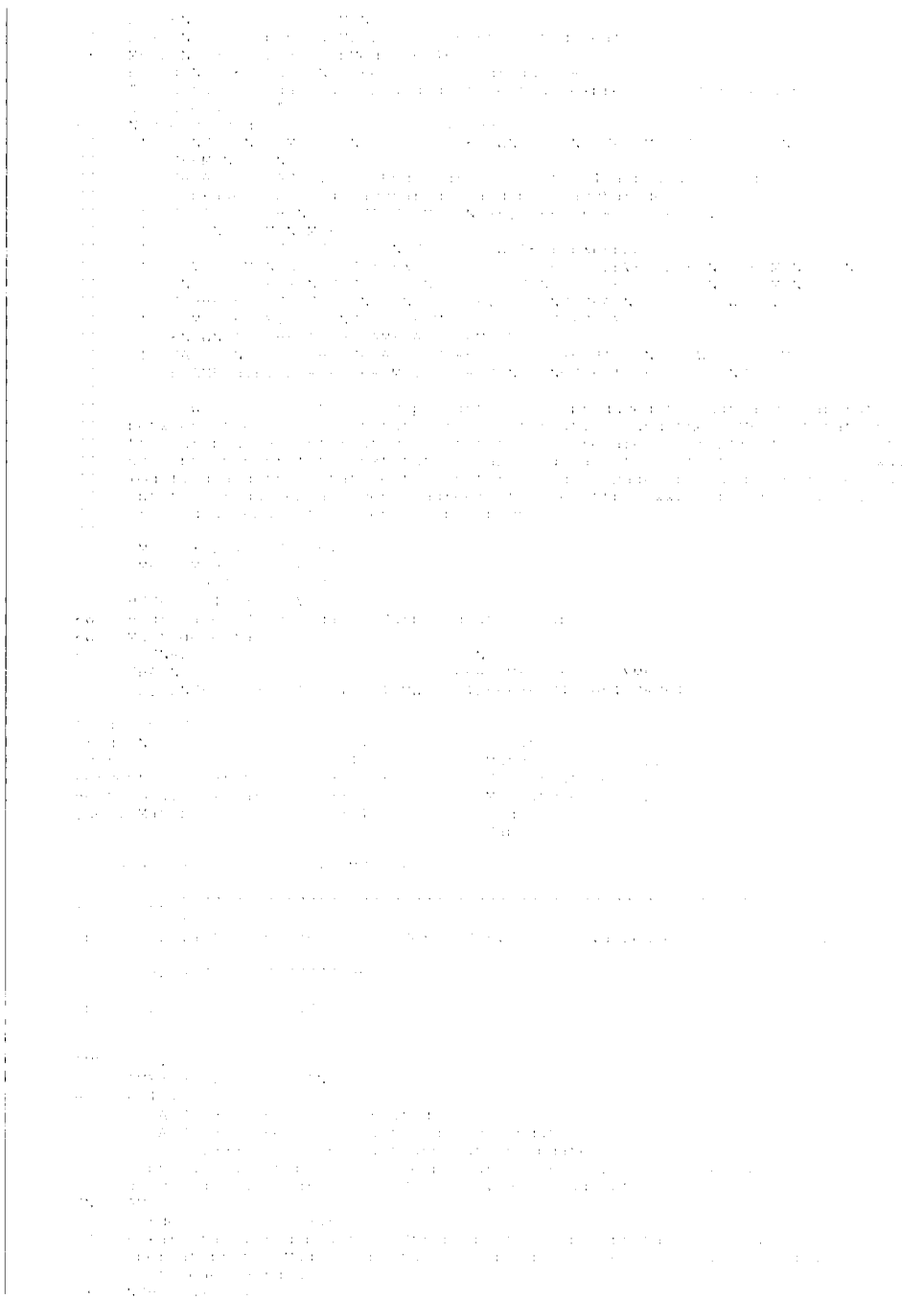
SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	117	65.7	428	1	AMY1_ORYSA	P17654 oryza sativ
2	104	58.4	383	1	AMC1_ORYSA	P27040 oryza sativ
3	71.5	40.2	368	1	AMY3_HORVU	P04747 hordeum vul
4	71.5	40.2	427	1	AMY2_HORVU	P04065 hordeum vul
5	69.5	39.0	429	1	AMY6_HORVU	P04750 hordeum vul
c 6	63	37.3	185	1	SLYD_HELPJ	Q9ZKB9 helicobacte
c 7	63	37.3	185	1	SLYD_HELPY	P56748 helicobacte
8	63	35.4	657	1	MMI5_MOUSE	Q54752 mus musculu
9	60	33.7	892	1	AIX7_HUMAN	Q75G55 homo sapien
c 10	59	34.9	289	1	E1A_ADE02	P03254 human aden
c 11	59	34.9	289	1	E1A_ADE05	P03255 human aden
c 12	58.5	34.6	2286	1	DPOE_HUMAN	Q07864 homo sapien
c 13	57.5	34.0	364	1	IF6B_PRVKA	P24827 pseudorabie
14	57	32.0	258	1	PRIA_LENED	Q01200 lentinula e
c 15	57	33.7	667	1	GCP4_HUMAN	Q96RJ1 homo sapien
16	56	31.5	317	1	TN11_HUMAN	Q14766 homo sapien
c 17	56	33.1	1104	1	G2D1_MOUSE	Q91157 mus musculu
18	55	30.9	626	1	RN17_MOUSE	Q99NG7 mus musculu
19	55	32.5	437	1	AMY8_HUMAN	P19531 homo sapien
20	55	32.5	437	1	AMY1_HUMAN	P04745 homo sapien
21	54.5	32.0	437	1	AMY3_HUMAN	P04746 homo sapien
22	54.5	32.0	437	1	AMY2_HUMAN	P04746 homo sapien
23	54.5	32.0	437	1	AMY2_MOUSE	P00687 mus musculu
c 24	54.5	32.0	640	1	SSE_MOUSE	Q9QXS6 mus musculu
25	54.5	32.0	640	1	SSE_HUMAN	Q15831 homo sapien
c 26	54	32.0	1154	1	PKD1_HUMAN	P98161 homo sapien
27	54	30.3	511	1	CYC1_MOUSE	P99028 mus musculu
28	54	30.3	1e3	1	Y...._[illegible]	[illegible]
29	54	30.3	52[illegible]	1	SIA7_HUMAN	Q92685 homo sapien
30	54	30.3	5[illegible]	1	MMP1_HUMAN	P03956 homo sapien
c 31	54	30.3	8[illegible]	1	YIP1_YEAST	P32844 saccharomyc
32	54	30.3	1455	1	PLC1_STRAU	Q03021 streptomyce
33	54	30.3	5401	1	HLPP_BPVIF	P16059 enterobacte
c 34	54.5	31.7	724	1	CAMP_MOUSE	P51829 mus musculu
c 35	54.5	31.7	6[illegible]	1	APP2_MOUSE	P51818 mus musculu
36	54	29.8	348	1	FLIG_MYCTU	P71869 mycobacteri
37	54	29.8	86[illegible]	1	GAT1_YEAST	P32329 saccharomyc
38	54	29.8	5[illegible]	1	VTCA_DEVIS	P24[illegible]
39	54	29.8	540[illegible]	1	[illegible]	[illegible]
c 40	54	30.3	1[illegible]	1	VNA_ENVEA	P[illegible]
41	54	29.8	1[illegible]	1	[illegible]	[illegible]
42	53.5	29.8	4[illegible]	1	[illegible]	[illegible]
c 43	53.5	31.1	5[illegible]	1	[illegible]	[illegible]
44	53.5	29.8	5[illegible]	1	RFA_BPVM2	P16[illegible]
45	53.5	29.8	1[illegible]	1	RS1_DROME	[illegible]

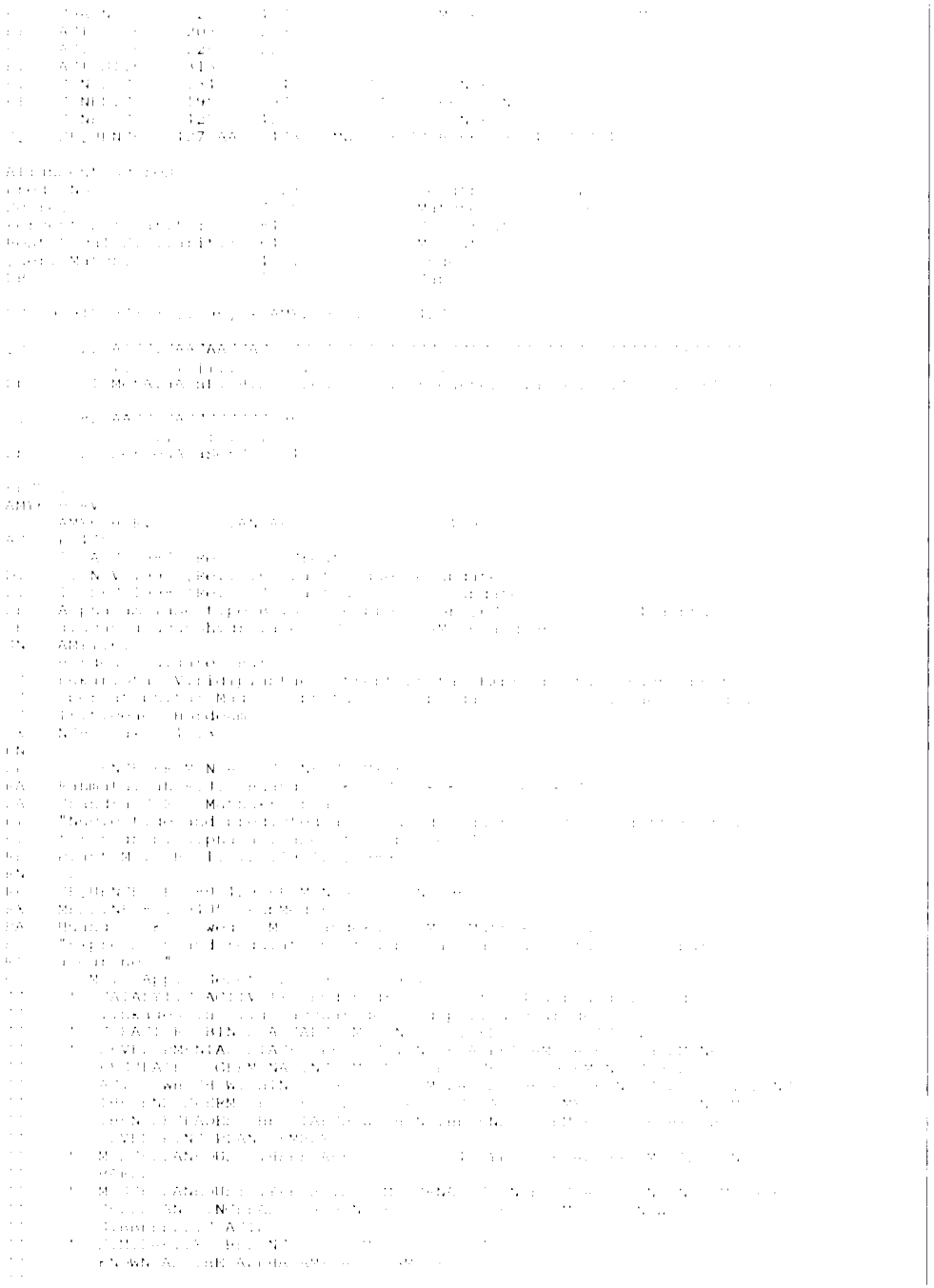
ALIGNMENTS

RESULT 1
AMY1_ORYSA
ID AMY1_ORYSA STANDARD; PRT; [illegible] AA.
AC P17654;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Alpha amylase precursor (EC [illegible]) (1,4-alpha-D-glucan
DE glucanohydrolase) (isozyme 1).
GN AMY1_1 OR AMY1A.
OS Oryza sativa (rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; [illegible]; Liliopsida; Poales;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Japonica;
RX MEDLINE=91264944; PubMed=[illegible];
RA Huang N., Sutliff T.D., Litts J.C., Rodriguez R.L.;
RT "Classification and characterization of the rice alpha-amylase
RT multigene family.";
RL Plant Mol. Biol. [illegible];
RN [2]
RP SEQUENCE FROM N.A. ([illegible]).
RC STRAIN=cv. Japonica;
RX MEDLINE=[illegible]; PubMed=[illegible];
RA [illegible];
RT "[illegible]";
RL [illegible].
CC -!- FUNCTION: IMPORTANT FOR BREAKDOWN OF ENDOSPERM STARCH
CC DURING GERMINATION.
CC -!- CATALYTIC ACTIVITY: Hydrolysis of [illegible]
CC linkages in oligosaccharides, but not polysaccharides.
CC -!- COFACTOR: BINDS A CALCIUM ION, WHICH IS ESSENTIAL FOR
CC -!- SUBUNIT: MONOMER.
CC -!- TISSUE SPECIFICITY: MOST ABUNDANT IN GERMINATING
CC CALLUS, YOUNG LEAVES.
CC -!- DEVELOPMENTAL STAGE: EXPRESSED AT A HIGH LEVEL [illegible]
CC IN THE ALEURONE LAYER CELLS OF MATURE SEEDS OF THE PLANT



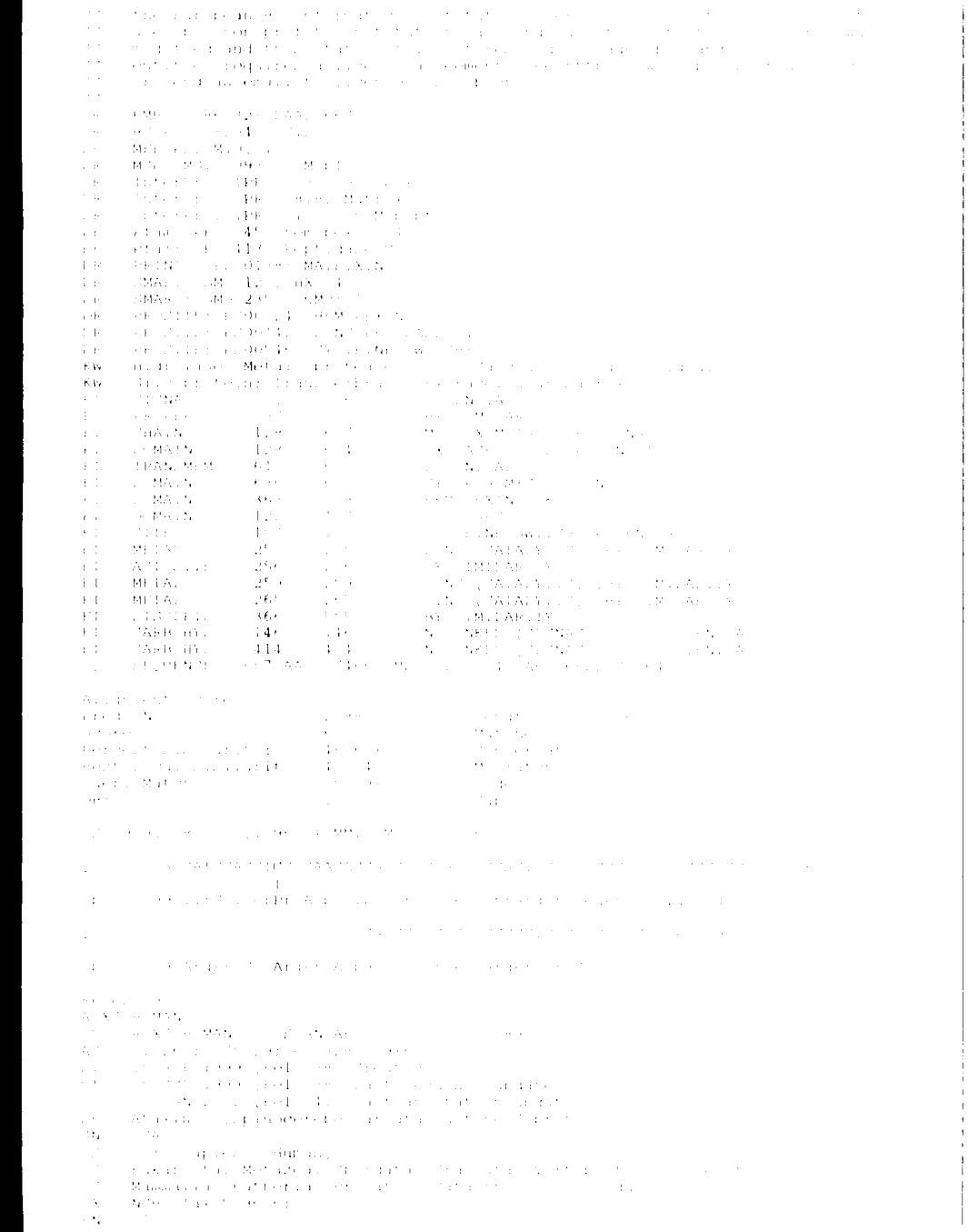


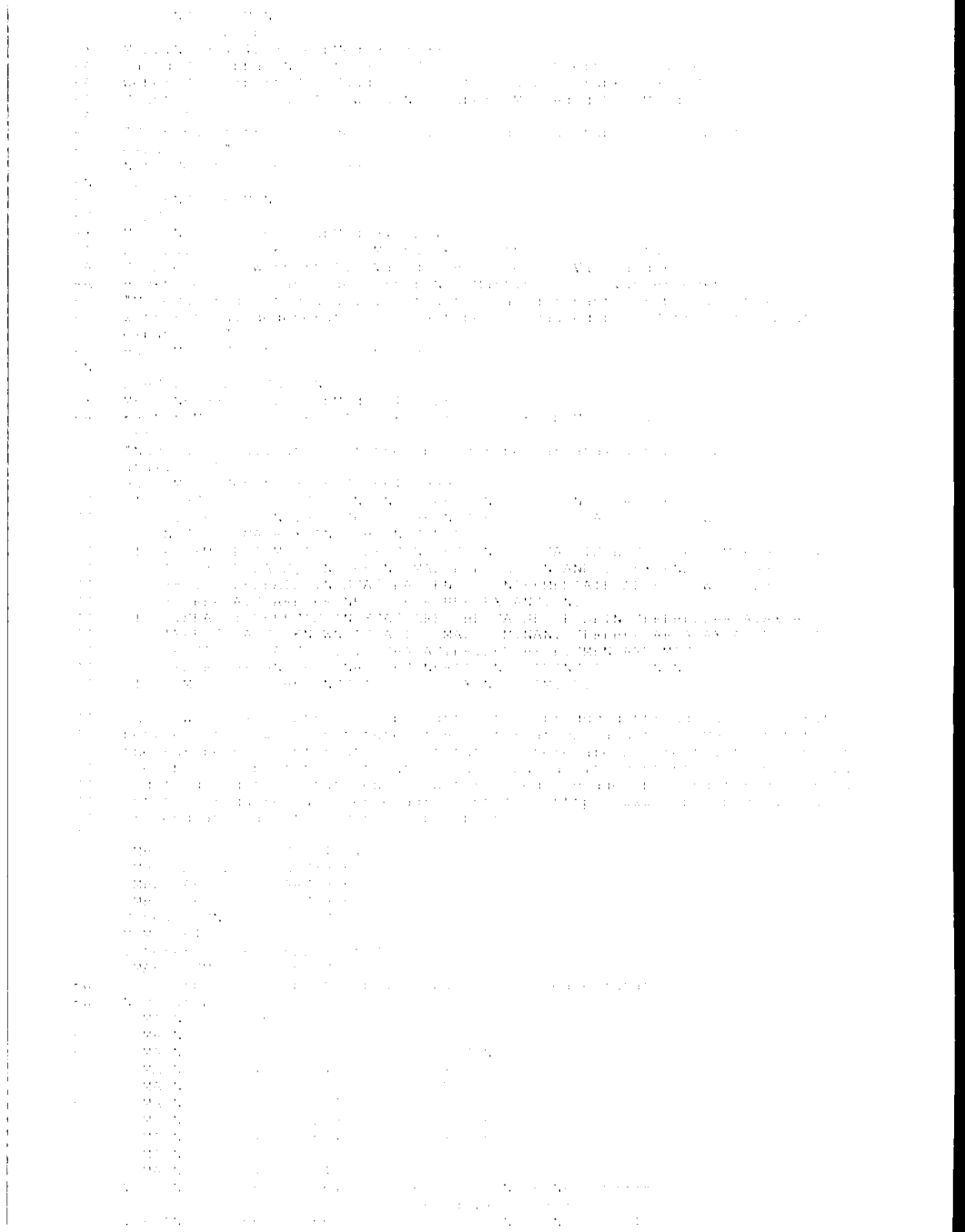
```
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Chandler P.M., Zwar J.A., Jacobsen J.V., Higgins T.J.V., Inglis A.S.;
RT   "The effects of gibberellic acid and abscisic acid on alpha-amylase
RT   mRNA levels in barley aleurone layers studies using an alpha-amylase
RT   cDNA clone.";
RL   Plant Mol. Biol. 3:407-418(1984).
CC   -!- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
CC       linkages in oligosaccharides and polysaccharides.
CC   -!- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
CC   -!- DEVELOPMENTAL STAGE: PRODUCTION OF ALPHA-AMYLASE IS HORMONALLY
CC       REGULATED. GERMINATING EMBRYOS PRODUCE THE HORMONE GIBBERELLIC
CC       ACID, WHICH WITHIN 10 HOURS STIMULATES THE ALEURONE CELLS COVERING
CC       THE ENDOSPERM OF THE SEED TO PRODUCE ALPHA-AMYLASE. THE ENZYME
CC       THEN DEGRADES THE STARCH WITHIN THE ENDOSPERM FOR USE BY THE
CC       DEVELOPING PLANT EMBRYO.
CC   -!- MISCELLANEOUS: THERE ARE AT LEAST 4 TYPES OF ALPHA AMYLASE IN
CC       BARLEY.
CC   -!- MISCELLANEOUS: TYPE B ISOZYME MRNA IS UNDETECTABLE IN UNSTIMULATED
CC       CELLS AND INCREASES A HUNDRED FOLD AFTER STIMULATION WITH
CC       GIBBERELLIC ACID.
CC   -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES. ALSO
CC       KNOWN AS THE ALPHA AMYLASE FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; K02638; AAA32933.1; -.
DR   PIR; A26267; A26267.
DR   HSSP; P04063; 1AVA.
DR   InterPro; IPR006461; Alpha_amylase.
DR   Pfam; PF00128; alpha-amylase; 1.
KW   Hydrolase; Glycosidase; Carbohydrate metabolism; Seed; Germination;
KW   Calcium; Multigene family; Signal.
FT   SIGNAL        1     24
FT   CHAIN        25    >368       ALPHA AMYLASE TYPE B ISOZYME.
FT   ACT_SITE    203     203       BY SIMILARITY.
FT   ACT_SITE    228     228       BY SIMILARITY.
FT   ACT_SITE    313     313       BY SIMILARITY.
FT   NON_TER     368     368
SQ   SEQUENCE   368 AA;  40797 MW;  A247EF5579[illegible] CRC64;

Alignment Scores:
Pred. No.:              0.233          Length:          368
Score:                  71.50          Matches:         1[illegible]
Percent Similarity:     84.00%         Conservative:    5
Best Local Similarity:  64.00%         Mismatches:      [illegible]
Query Match:            40.17%         Indels:          1
DB:                     1              Gaps:            1

US-09-915-873-3 (1-96) x AMY3_HORVU (1-368)

Qy       22 ATGGTCAACAAGCACTTGCTCTCCCTGTCCGTGCTCATC[illegible] 81
               |||   ||||||||   ||||||||||||:::   ::::|| ||||||||  ||:::
Db        1 MetAlaAsnLysHis---LeuSerLeuSerLeuPheLeuValLeuLeuGlyLeuSerAla 19

Qy       82 AACCTCACCGCCGGC 96
               ::::|||   ::::|||
Db       20 SerLeuAlaSerGly 24

RESULT 4
AMY2_HORVU
ID   AMY2_HORVU     STANDARD;      PRT;   427 AA.
AC   P04063;
DT   01-NOV-1986 (Rel. 03, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Alpha-amylase type B isozyme precursor (EC 3.2.1.1) (1,4-alpha-D-
DE   glucan glucanohydrolase) (AMY[illegible]) [illegible] alpha-amylase).
GN   AMY1.2.
OS   Hordeum vulgare (Barley).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; [illegible]
OC   Triticeae; Hordeum.
OX   NCBI_TaxID=4513;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Rahmatullah R.J., [illegible]
RA   Chandra G.R., Muthukrishnan S.;
RT   "Nucleotide and predicted amino acid sequences of two different genes
RT   for high-pI alpha amylases from barley.";
RL   Plant Mol. Biol. [illegible]
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE [illegible]; PubMed [illegible]
RA   Rogers J.C.;
RT   "Two barley alpha-amylase [illegible]
RT   aleurone cells.";
RL   J. Biol. Chem. [illegible]
RN   [3]
RP   X-RAY CRYSTALLOGRAPHY [illegible]
RX   MEDLINE 9425[illegible]; PubMed [illegible]
RA   Kadziola A., Abe J.[illegible]
RT   "Crystal and molecular structure of barley alpha-amylase.";
RL   J. Mol. Biol. [illegible]
RN   [4]
RP   X-RAY CRYSTALLOGRAPHY [illegible] COMPLEX WITH [illegible]
RC   STRAIN=cv. Menuet;
RX   MEDLINE 98298441; PubMed [illegible]
RA   Vallee F., Kadziola A., [illegible] K.W.
RA   Svensson B., Haser R.;
RT   "Barley alpha-amylase [illegible]
RT   crystal structure of the complex [illegible] A resolution.";
RL   Structure [illegible]
CC   -!- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
CC       linkages in oligosaccharides and polysaccharides.
CC   -!- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
CC   -!- DEVELOPMENTAL STAGE: PRODUCTION OF ALPHA-AMYLASE IS HORMONALLY
CC       REGULATED. GERMINATING EMBRYOS PRODUCE THE HORMONE GIBBERELLIC
CC       ACID, WHICH WITHIN [illegible] THE ALEURONE CELLS COVERING
CC       THE ENDOSPERM OF THE SEED TO PRODUCE ALPHA-AMYLASE. THE ENZYME
CC       THEN DEGRADES THE STARCH WITHIN THE ENDOSPERM FOR USE BY THE
CC       DEVELOPING PLANT EMBRYO.
CC   -!- INDUCTION: TYPE B ISOZYME MRNA IS UNDETECTABLE IN UNSTIMULATED
CC       CELLS AND INCREASES A HUNDRED [illegible] AFTER STIMULATION WITH
CC       GIBBERELLIC ACID.
CC   -!- MISCELLANEOUS: THERE ARE AT LEAST 4 TYPES OF ALPHA AMYLASE IN
CC       BARLEY.
CC   -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES. ALSO
CC       KNOWN AS THE ALPHA AMYLASE FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X15226; CAA33298.1; -.
DR   EMBL; K02637; AAA[illegible]; -.
DR   PIR; A00847; ALBHB.
DR   PIR; [illegible]
DR   PDB; 1AMY; 13-MAY-98.
DR   PDB; 1AVA; 1[illegible]-MAR-[illegible].
DR   InterPro; IPR006461; Alpha_amylase.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   PRINTS; PR00110; ALPHAAMYLASE.
KW   Hydrolase; Glycosidase; Carbohydrate metabolism; Seed; [illegible]
KW   Calcium; Multigene family; Signal; 3D-structure.
FT   SIGNAL        1     24
```



```
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch)
CC   -----------------------------------------------------------------------
DR   EMBL; AE001533; AAD06631.1; -.
DR   InterPro; IPR001179; FKBP_PPIase.
DR   Pfam; PF00254; FKBP; 1.
DR   PROSITE; PS00453; FKBP_PPIASE_1; FALSE_NEG.
DR   PROSITE; PS00454; FKBP_PPIASE_2; FALSE_NEG.
DR   PROSITE; PS50059; FKBP_PPIASE_3; 1.
KW   Isomerase; Rotamase; Complete proteome.
FT   DOMAIN        1     99       PPIASE, FKBP-TYPE.
FT   DOMAIN      158    185       GLY-RICH.
SQ   SEQUENCE   185 AA;  20100 MW;  2B48E882AC450B0CB CRC64;

Alignment Scores:
Pred. No.:                2.7          Length:          185
Score:                    63.00        Matches:         12
Percent Similarity:       54.17%       Conservative:    1
Best Local Similarity:    50.00%       Mismatches:      9
Query Match:              37.28%       Indels:          2
DB:                       1            Gaps:            1

US-09-915-873-3 (1-66) x SLYD_HELPY (1-185)

Qy      85 CCTTCCTCCTCAGCCCGAGCAGCATCAGCAGCGACAGCAGTCAGAAAGTCTTTCTTCA 2
           ||||||||||          ||||||      |||:::   |||   |||       |||
Db     162 GlyCysCysGlyCl,HisGlyCl,HisCl,Cl,C,Cl,Cl,Cl,Cl,Cl,Cl,Cl,Cl,   1

Qy      25 CCATCGTGTTCA 14
           ||||||||||
Db     181 ---SerCysSer 183

RESULT 7
SLYD_HELPY
ID   SLYD_HELPY     STANDARD;      PRT;   185 AA.
AC   O25748;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   FKBP-type peptidyl-prolyl cis-trans isomerase slyD (EC 5.2.1.8)
DE   (PPIase) (Rotamase).
GN   SLYD OR HP1123.
OS   Helicobacter pylori (Campylobacter pylori).
OC   Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC   Helicobacter.
OX   NCBI_TaxID=210;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN 26695 / ATCC 700392;
RX   MEDLINE=97394467; PubMed=9252185;
RA   Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
RA   Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA   Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA   Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA   McKenney K., FitzGerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA   Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA   Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,
RA   Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA   Venter J.C.;
RT   "The complete genome sequence of the gastric pathogen Helicobacter
RT   pylori.";
RL   Nature 388:539-547(1997).
CC   -!- FUNCTION: PPIASES ACCELERATE THE FOLDING OF PROTEINS (BY
CC       SIMILARITY).
CC   -!- CATALYTIC ACTIVITY: CIS-TRANS ISOMERIZATION OF PROLINE IMIDIC
CC       PEPTIDE BONDS IN OLIGOPEPTIDES.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC   -!- SIMILARITY: BELONGS TO THE FKBP-TYPE PPIASE FAMILY.
```


```
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch)
CC   -----------------------------------------------------------------------
DR   EMBL; AE001533; AAD06631.1; -.
DR   InterPro; IPR001179; FKBP_PPIase.
DR   Pfam; PF00254; FKBP; 1.
DR   PROSITE; PS00453; FKBP_PPIASE_1; FALSE_NEG.
DR   PROSITE; PS00454; FKBP_PPIASE_2; FALSE_NEG.
DR   PROSITE; PS50059; FKBP_PPIASE_3; 1.
KW   Isomerase; Rotamase; Complete proteome.
FT   DOMAIN        1     99       PPIASE, FKBP-TYPE.
FT   DOMAIN      158    185       GLY-RICH.
SQ   SEQUENCE   185 AA;  20100 MW;  2B48E882AC450B0CB CRC64;

Alignment Scores:
Pred. No.:                2.7          Length:          185
Score:                    63.00        Matches:         12
Percent Similarity:       54.17%       Conservative:    1
Best Local Similarity:    50.00%       Mismatches:      9
Query Match:              37.28%       Indels:          2
DB:                       1            Gaps:            1

US-09-915-873-3 (1-66) x SLYD_HELPY (1-185)

Qy      85 CCTTCCTCCTCAGCCCGAGCAGCATCAGCAGCGACAGCAGTCAGAAAGTCTTTCTTCA 2
           ||||||||||          ||||||      |||:::   |||   |||       |||
Db     162 GlyCysCysGlyCl,HisGlyCl,HisCl,Cl,C,Cl,Cl,Cl,Cl,Cl,Cl,Cl,Cl,   1

Qy      25 CCATCGTGTTCA 14
           ||||||||||
Db     181 ---SerCysSer 183

RESULT 7
SLYD_HELPY
ID   SLYD_HELPY     STANDARD;      PRT;   185 AA.
AC   O25748;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   FKBP-type peptidyl-prolyl cis-trans isomerase slyD (EC 5.2.1.8)
DE   (PPIase) (Rotamase).
GN   SLYD OR HP1123.
OS   Helicobacter pylori (Campylobacter pylori).
OC   Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC   Helicobacter.
OX   NCBI_TaxID=210;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN 26695 / ATCC 700392;
RX   MEDLINE=97394467; PubMed=9252185;
RA   Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
RA   Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA   Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA   Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA   McKenney K., FitzGerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA   Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA   Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,
RA   Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA   Venter J.C.;
RT   "The complete genome sequence of the gastric pathogen Helicobacter
RT   pylori.";
RL   Nature 388:539-547(1997).
CC   -!- FUNCTION: PPIASES ACCELERATE THE FOLDING OF PROTEINS (BY
CC       SIMILARITY).
CC   -!- CATALYTIC ACTIVITY: CIS-TRANS ISOMERIZATION OF PROLINE IMIDIC
CC       PEPTIDE BONDS IN OLIGOPEPTIDES.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC   -!- SIMILARITY: BELONGS TO THE FKBP-TYPE PPIASE FAMILY.
```

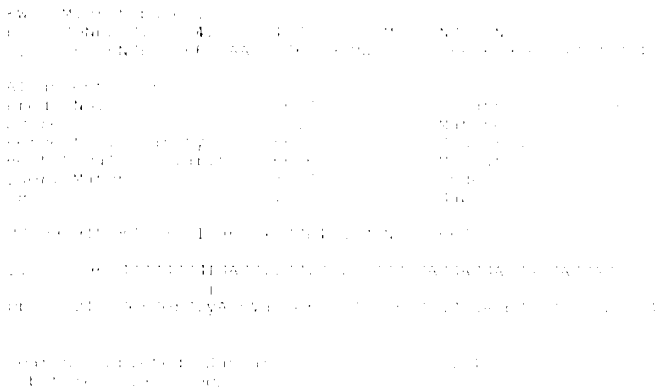




```
FT                              /FTId=VAR_011824.
FT   CONFLICT     105    105       P -> H (IN REF. 2; AAC14164).
FT   CONFLICT     129    129       G -> S (IN REF. 2; AAC14164).
FT   CONFLICT     862    862       V -> M (IN REF. 2).
FT   CONFLICT     888    892       PKARP -> VGNGL (IN REF. 2.
FT                              AAC39765/AAC14163).
SQ   SEQUENCE   892 AA;  95451 MW;  9AEA7671772B4C9F CRC64;

Alignment Scores:
Pred. No.:                7.04          Length:          892
Score:                   60.00          Matches:          18
Percent Similarity:      59.26%         Conservative:      3
Best Local Similarity:   48.15%         Mismatches:        9
Query Match:             33.71%         Indels:            2
DB:                      1              Gaps:              2

US-09-915-873-3 (1-96) x AIX7_HUMAN (1-892)

Qy      18 CACGATCGTCAACAAGCACTTCCTCTCGCTGTCGTC  53
           ||| |||  |||  |||  |||::|||  ||||||  |||||||
Db     401 HisProAspSerGlnGlnProProGlnProLeuArgAspProHisProAlaProProArg 420

Qy      72 GCTGAGCAGCAACCTCACCGC 92
           |||::||||||:::
Db     421 IhrSerGlnGluProHisGln 427

RESULT 10
E1A_ADEu2
ID   E1A_ADE02      STANDARD;      PRT;   289 AA.
AC   P03254;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Early E1A 32 kDa protein [Contains: Early E1A 29 kDa protein].
OS   Human adenovirus type 2.
OC   Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX   NCBI_TaxID=10515;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE 91012104; PubMed 551296.
RA   Perricaudet M., Akusjaervi G., Virtanen A., Pettersson U.;
RT   "Structure of two spliced mRNAs from the transforming region of human
RT   subgroup C adenoviruses.";
RL   Nature 281:694-696(1979).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=83056843; PubMed 7142161;
RA   Cingeras T.R., Sciaky D., Gelinas R.E., Bing-Dong J., Yen C.E.,
RA   Kelly M.M., Bullock P.A., Parsons B.L., O'Neill K.E., Roberts R.J.,
RT   "Nucleotide sequences from the adenovirus-2 genome.";
RL   J. Biol. Chem. 257:13475-13491(1982).
CC   -!- FUNCTION: TRANS-ACTIVATES EARLY VIRAL PROMOTERS AND S-MU CELLULAR
CC       PROMOTERS. THE TUMORIGENESIS-RESTRAINING ACTIVITY OF E1A MAY BE
CC       RELATED TO THE DISRUPTION OF THE HUMAN CIBP 21P ZINC-INDEX TRANSCRIPTION
CC       THE CIBP BINDING MOTIF.
CC   -!- ALTERNATIVE PRODUCTS: THESE PROTEINS ARE PRODUCTS OF TWO DIFFERENT
CC       MRNAS SPLICED FROM THE SAME PRIMARY TRANSCRIPT FROM THE E1 REGION
CC       OF THE GENOME.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaborati
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announ
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; J01917; AAA92199.1; -.
DR   PIR; A03824; Q2AD2.
DR   TRANSFAC; T00209; -.
DR   InterPro; IPR003854; Adeno_E1A.
DR   Pfam; PF02703; Adeno_E1A; 1.
KW   Transcription regulation; Activator; Early protein.
```

```
KW   Alternative splicing; DNA-binding; Zinc-finger; DNA replication.
FT   ZN_FING      205    227       C2H2 TYPE.
FT   SITE         209    210       CIBP-BINDING MOTIF.
FT   VARSPLIC     116    147       MISSING (IN ISOFORM E1A-289R).
FT   CONFLICT      98     98       L -> R (IN REF. 2).
FT   CONFLICT      81     81       L -> R (IN REF. 2).
SQ   SEQUENCE   289 AA;  [illegible] MW;  [illegible] CRC64;

Alignment Scores:
Pred. No.:                [illegible]    Length:          289
Score:                    [illegible]    Matches:          12
Percent Similarity:       [illegible]    Conservative:      1
Best Local Similarity:    41.11%         Mismatches:        5
Query Match:              24.72%         Indels:            0
DB:                       1              Gaps:              0

US-09-915-873-3 (1-96) x E1A_ADEu2 (1-289)

Qy      41 GCTGCGAGCGGTGTTCAGCCGTCCTCTGGATGATCTGTTGCAGTACGTGTCTCCAGTG  100
           ||| |||  |||  |||  |||::|||  ||||||  |||||||  |||  |||  |||
Db     127 AlaAlaProProGlnProAspAspLeuGluMetCysTyrValSerProVal  179

Qy     101 GAAGACAAAGCCGGCGAAGGTATTAAAGAGGAGATGGAAGAC  140
           |||::||||||:::
Db     180 GluAspLysAlaGlyGluGlyIleLysGluGluMetGluAsp  221

RESULT 11
E1A_ADE05
ID   E1A_ADE05      STANDARD;      PRT;   289 AA.
AC   P03255; Q64825; Q64826;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Early E1A 32 kDa protein [Contains: Early E1A 29 kDa protein].
OS   Human adenovirus type 5.
OC   Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX   NCBI_TaxID=28285;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE 81165557; PubMed 6259158;
RA   Van Ormondt H., Maat J., Van Beveren C.P.;
RT   "The nucleotide sequence of the transforming gene, region 1
RT   adenovirus type 5 DNA.";
RL   Gene 11:299-308(1980).
RN   [2]
RP   MUTAGENESIS.
RX   MEDLINE 92252269; PubMed [illegible];
RA   Webster [illegible], Chen R., Mack D., Accola L., Olsen D.S., Horwitz M.,
RA   Wu P.Y.-S., Fitzgerald R.J.;
RT   "Conversion of the E1A 12S1 zinc finger to a nonfunctional zinc
RT   finger by a single point mutation.";
RL   Proc. Natl. Acad. Sci. U.S.A. [illegible](1992).
RN   [3]
RP   COMPLETE GENOME.
RX   MEDLINE 93308745; PubMed [illegible];
RA   Chroboczek J., Bieber F., Jacrot B.;
RT   "The sequence of the genome of adenovirus type 5 and its
RT   with the genome of adenovirus type 2.";
RL   Virology 186:280-285(1992).
CC   -!- FUNCTION: TRANS-ACTIVATES EARLY VIRAL PROMOTERS AND S-MU CELLULAR
CC       PROMOTERS. THE TUMORIGENESIS-RESTRAINING ACTIVITY OF E1A MAY BE
CC       RELATED TO THE DISRUPTION OF THE HUMAN CIBP 21P ZINC-INDEX TRANSCRIPTION
CC       THE CIBP BINDING MOTIF.
CC   -!- ALTERNATIVE PRODUCTS: THESE PROTEINS ARE PRODUCTS OF TWO DIFFERENT
CC       MRNAS SPLICED FROM THE SAME PRIMARY TRANSCRIPT FROM THE E1 REGION
CC       OF THE GENOME.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaborati
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announ
```

```
CC   This SWISS-PROT entry is copyright. It is produced through a collaborative
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------- --
DR   EMBL; D00633; -; NOT_ANNOTATED_CDS.
DR   PIR; A36255; EDBESP.
DR   InterPro; IPR003403; IE68.
DR   Pfam; PF02479; IE68; 1.
KW   Early protein.
FT   DOMAIN       215    344       ASP/GLU RICH (ACIDIC).
SQ   SEQUENCE   364 AA;  39613 MW;  CE1329ECD7E1515 CRC64;

Alignment Scores:
Pred. No.:                14              Length:          364
Score:                    57.50           Matches:         34
Percent Similarity:       70.00%          Conservative:    1
Best Local Similarity:    65.00%          Mismatches:      5
Query Match:              34.02%          Indels:          1
DB:                       1               Gaps:            1

US-09-915-873-3 (1-96) x IE68_EBVKA (1-364)

Qy      93 GGCGGTGAGGTTGCTGCTCAGCCCGAGGAGGACC--GATGAGGAVGAGAVGAGAVGAGAAA 57
           ||||||||||                 ||||||||   |||||||||||:|:::|:|.|||
Db     325 GlyGlyGluAspGlyGluAspGl,GluGluAspGluAspGluAspGlyGluGlyGluGlu 411

RESULT 14
PRIA_LENED
ID   PRIA_LENED     STANDARD;      PRT;   258 AA.
AC   Q01200;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   01-JUL-1993 (Rel. 26, Last annotation update)
DE   PriA protein precursor.
GN   PRIA.
OS   Lentinula edodes (Shiitake mushroom) (Lentinus edodes).
OC   Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
OC   Agaricales; Tricholomataceae; Lentinula.
OX   NCBI_TaxID=5353;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FMC2;
RX   MEDLINE 92290272; PubMed 1601301;
RA   Kajiwara S., Yamaoka K., Hori K., Miyazawa H., Saito T., Kanda T.,
RA   Shishido K.;
RT   "Isolation and sequence of a developmentally regulated protein kinase
RT   gene, priA, from the basidiomycete Lentinus edodes.";
RL   Gene 114:173-178(1992).
CC   -!- FUNCTION: MAY PLAY A ROLE DURING THE BEGINNING OF FRUITING.
CC   -!- DEVELOPMENTAL STAGE: DEVELOPMENTALLY REGULATED WITH PRIMORDIA/
CC       IMMATURE FRUITING BODIES HAVING MUCH HIGHER LEVELS OF PRIA.
CC   ---------------------------------------------------------------------- --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------- --
DR   EMBL; X60956; CAA43289.1; -.
DR   PIR; S23106; S23106.
KW   Signal.
FT   SIGNAL        1     24       POTENTIAL.
FT   CHAIN        24    258       PRIA PROTEIN.
FT   DOMAIN       41    242       CYS-RICH.
SQ   SEQUENCE   258 AA;  26725 MW;  247141B79BAB2C0B CRC64;

Alignment Scores:
Pred. No.:                [illegible]     Length:          [illegible]
Score:                    [illegible]     Matches:         [illegible]
Percent Similarity:       [illegible]     Conservative:    [illegible]
Best Local Similarity:    [illegible]     Mismatches:      [illegible]
Query Match:              [illegible]     Indels:          [illegible]
DB:                       [illegible]     Gaps:            [illegible]

US-09-915-873-3 (1-96) x PriA_LENED (1-258)

Qy      [illegible]
           [illegible]
Db      [illegible]

RESULT 15
[illegible]
ID   [illegible]     STANDARD;      [illegible]   [illegible] AA.
AC   [illegible];
DT   [illegible] (Rel. 41, Created)
DT   [illegible] (Rel. 41, Last sequence update)
DT   [illegible] (Rel. 41, Last annotation update)
DE   Gamma-tubulin complex component 4 (GCP-4) (hGCP4) (GCP4).
GN   GCP4 OR ZGC.
OS   Homo sapiens (human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL PROTEIN SEQUENCE.
RC   TISSUE=Neuroblastoma;
RX   MEDLINE 20029771; PubMed [illegible];
RA   Hale T., Raymond Bessling D., Gould K.L., Marcus S., Balasubramanian M.;
RA   Gilliland G.L., Carter D., Lee E., Edwards R., Ward M.;
RT   "Human [illegible] a novel protein associated with [illegible] tubulin-like
RT   protein family.";
RL   J. Cell Biol. 147:959-968(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RA   Strausberg R.;
RL   Submitted (AUG 2001) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF [illegible] FROM N.A.
RC   TISSUE=Teratocarcinoma;
RA   Imanaka T., Ota T., Morishita R., Tanimura T., Saito K., Ohara O.,
RA   Nishikawa T., Nomura N., Isogai T., Asakawa S., Kimura J., Shibuya K.,
RA   Matsuzawa H., Ishii S., Kikuno R., Suzuki K., Yamamoto J., Wakamatsu A.,
RA   Nakamura Y., Nagahari K., Misawa S., Tsuchiya N.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Gamma tubulin complex is composed of several [illegible]
CC       GCP5, GCP4, GCP5 and GCP6.
CC   -!- SUBCELLULAR LOCATION: [illegible].
CC   -!- TISSUE SPECIFICITY: Brain, expressed.
CC   -!- SIMILARITY: BELONGS TO [illegible] FAMILY.
CC   ---------------------------------------------------------------------- --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------- --
DR   EMBL; AJ249677; CAB62[illegible];
DR   EMBL; BC[illegible]; AAH[illegible];
DR   EMBL; BC012[illegible]; AAH[illegible];
DR   EMBL; AK[illegible]; BAB[illegible];
```



GenCore version 5.1.x

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 17, 2003, 09:59:27 ; Search time 23 Seconds
(without alignments)
1720.045 Million cell updates/sec

Title: US-09-915-873-3
Perfect score: 178
Sequence: 1 accatgcaggtcctgaacac.....ncaacaacctcgccgccgcgac 96

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 134316

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO_sp_1/tmp/STN16/5_out_14012003_171622_17295/app_query.fasta_1.seq
-DB=SPTREMBL_21 -QFMT=fasta -SUFFIX=n2p.rspt -MINMATCH=0.1 -MAXMATCH=1
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USPR=US09915873_accgn_1_1_19_rumat_14012003_171622_17295 -NOQS=1 -TOP=1
-NO_XLPXY -NO_MMAP -IAPGPQUERY -NEW_SCORES=0 -WAIT -U=NONE -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAP=6.0 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SPTREMBL_21:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

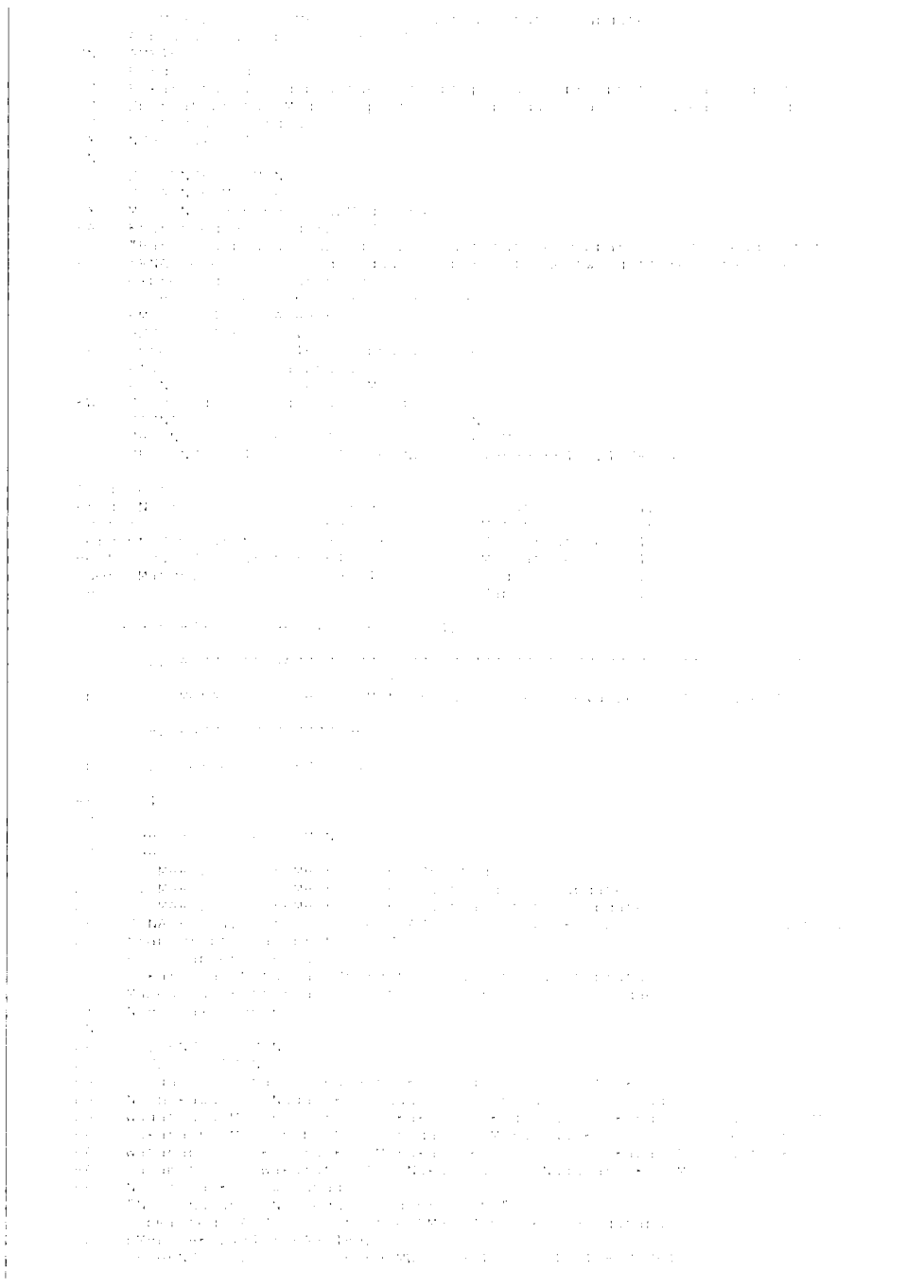
Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

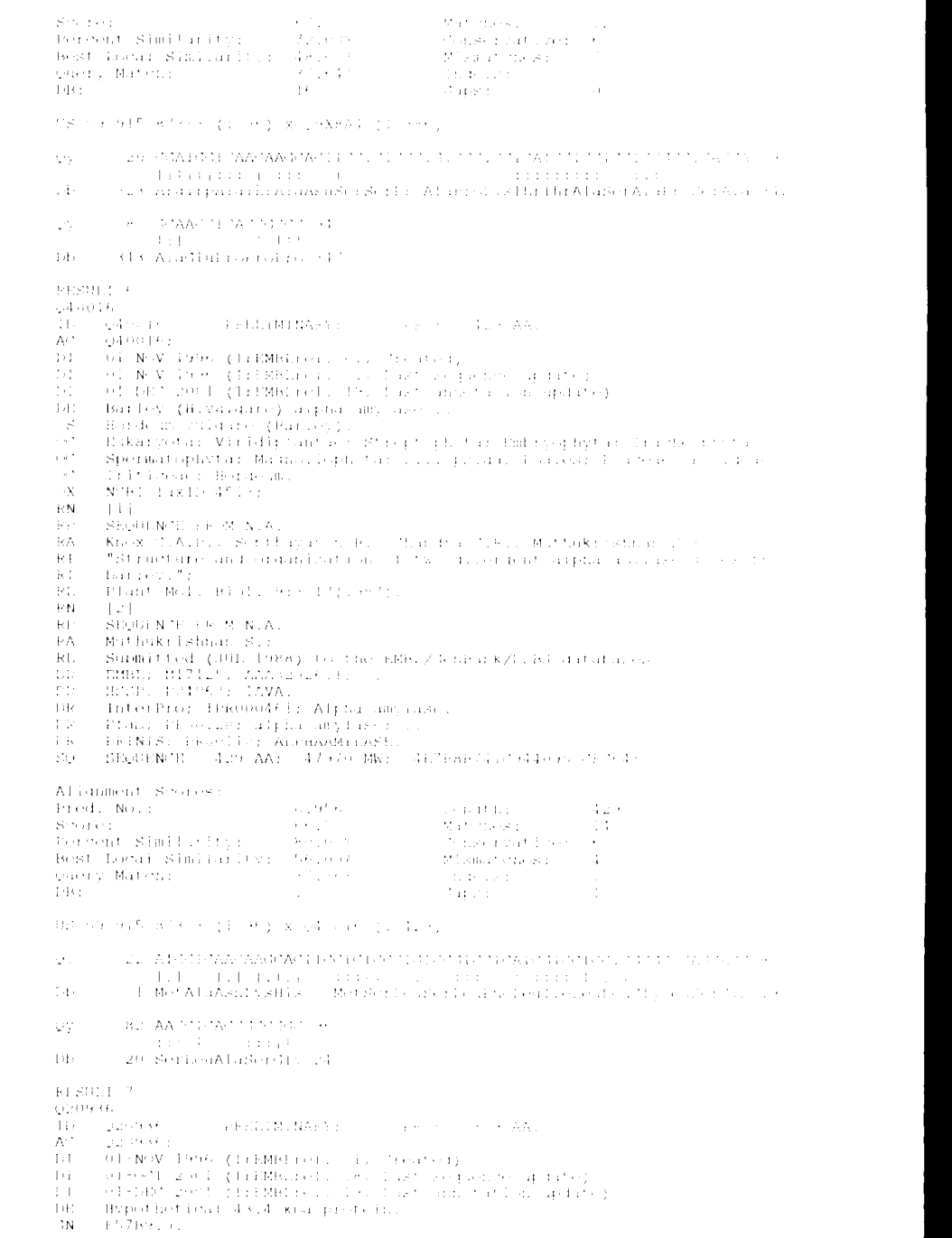
Result No.	Score	Query Match	Length	DB	ID	Description
1	104	58.4	424	10	Q94DK9	Q94DK9 Oryza sativa
2	71.6	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
3	64.5	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
4	62	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
5	62	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
6	60.5	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
7	60	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
8	60	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
9	60	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
10	72.5	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
11	70	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
12	70	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
13	70	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
14	69.5	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
15	68	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
16	68	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
17	68	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
18	68	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
19	68	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
20	68	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
21	67.5	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
22	67.5	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
23	67.5	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
24	67.5	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
25	67.5	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
26	67.5	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
27	67	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
28	67	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
29	67	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
30	66.5	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
31	66.5	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
32	66.5	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
33	66.5	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
34	66.5	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
35	66.5	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
36	66.5	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
37	66	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
38	66	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
39	66	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
40	66	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
41	66	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
42	66	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
43	65.5	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
44	65.5	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
45	65.5	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

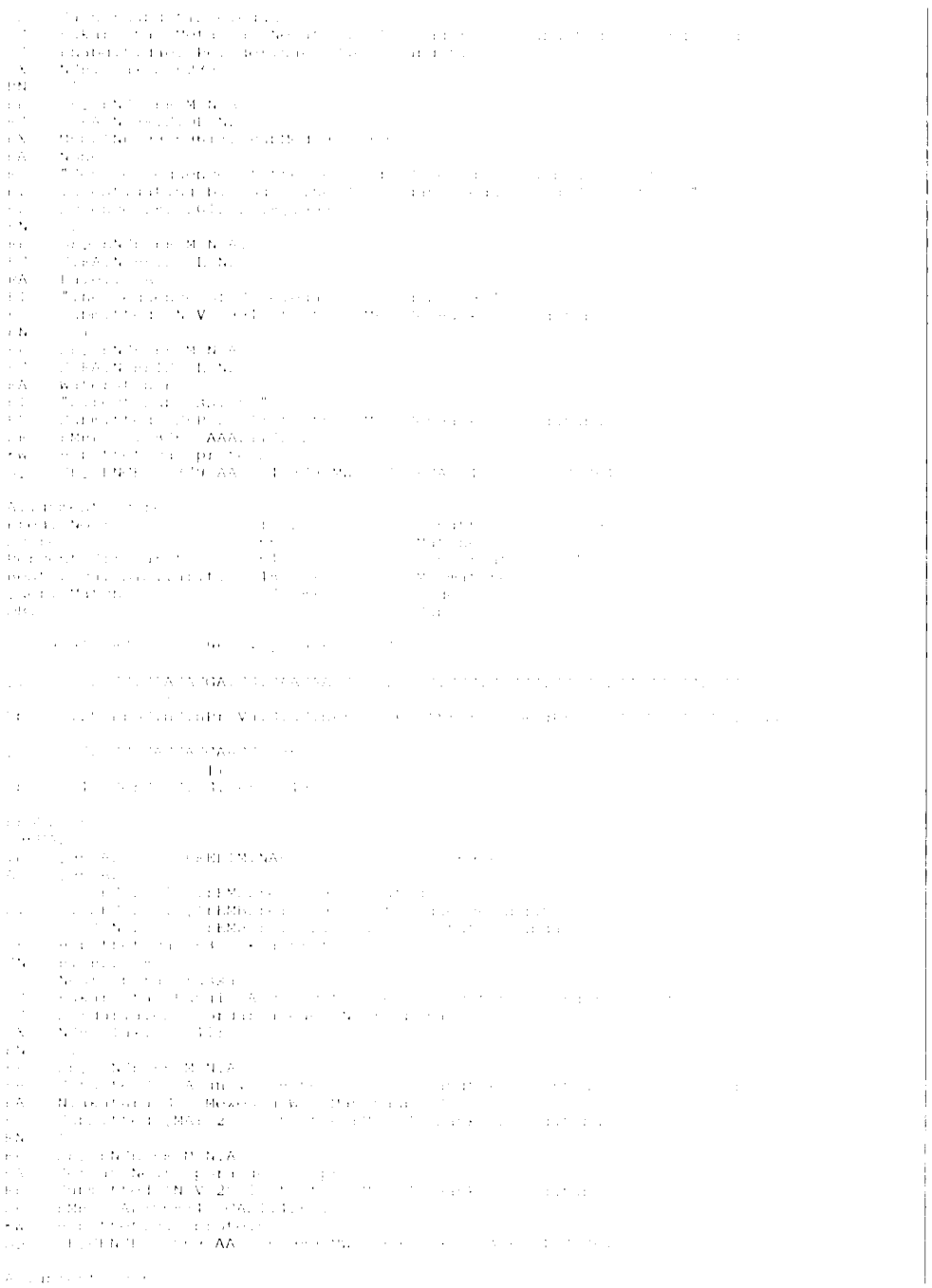
ALIGNMENTS

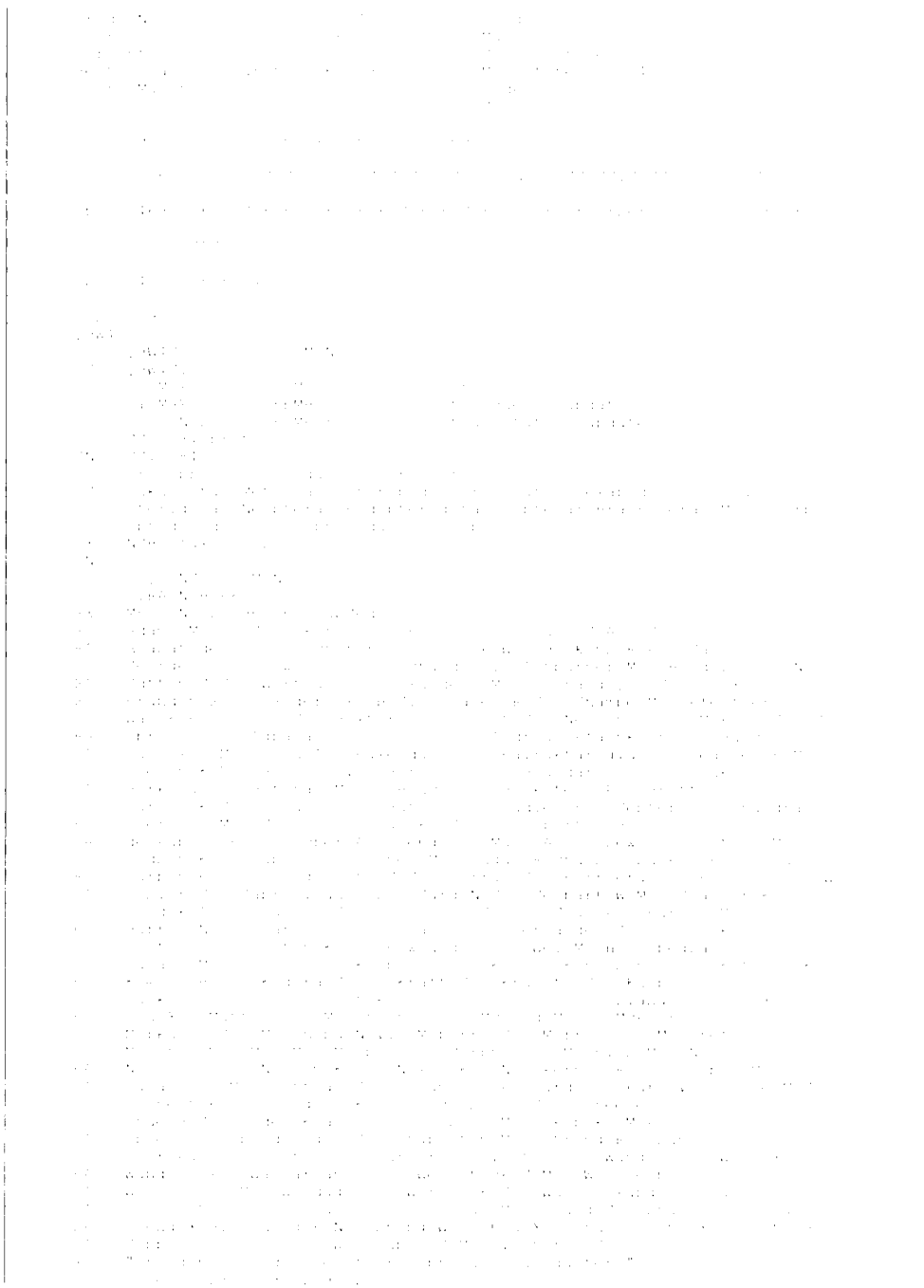
RESULT 1
Q94DK9
ID Q94DK9 PRELIMINARY; PRT; 424 AA.
AC Q94DK9;
DT 01-JAN-2001 (TREMBLrel. 19, Created)
DT 01-JAN-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 19, Last annotation update)
DE Putative alpha-hydroxylase.
GN P0514H03.12.
OS Oryza sativa (japonica).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Oryza.
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K., Kikuchi S.;
RT "Oryza sativa Nipponbare (WA, BAC, BIBAC, NA, and RA) genomic DNA.";
RT clone:P0514H03.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AP003275; BAB69610.1; -.
DR InterPro; IPR000461; Alpha_oxygenase.
DR Pfam; PF00248; alpha_oxygenase; 1.
SQ SEQUENCE 424 AA; 47662 MW; [illegible] CRC64;



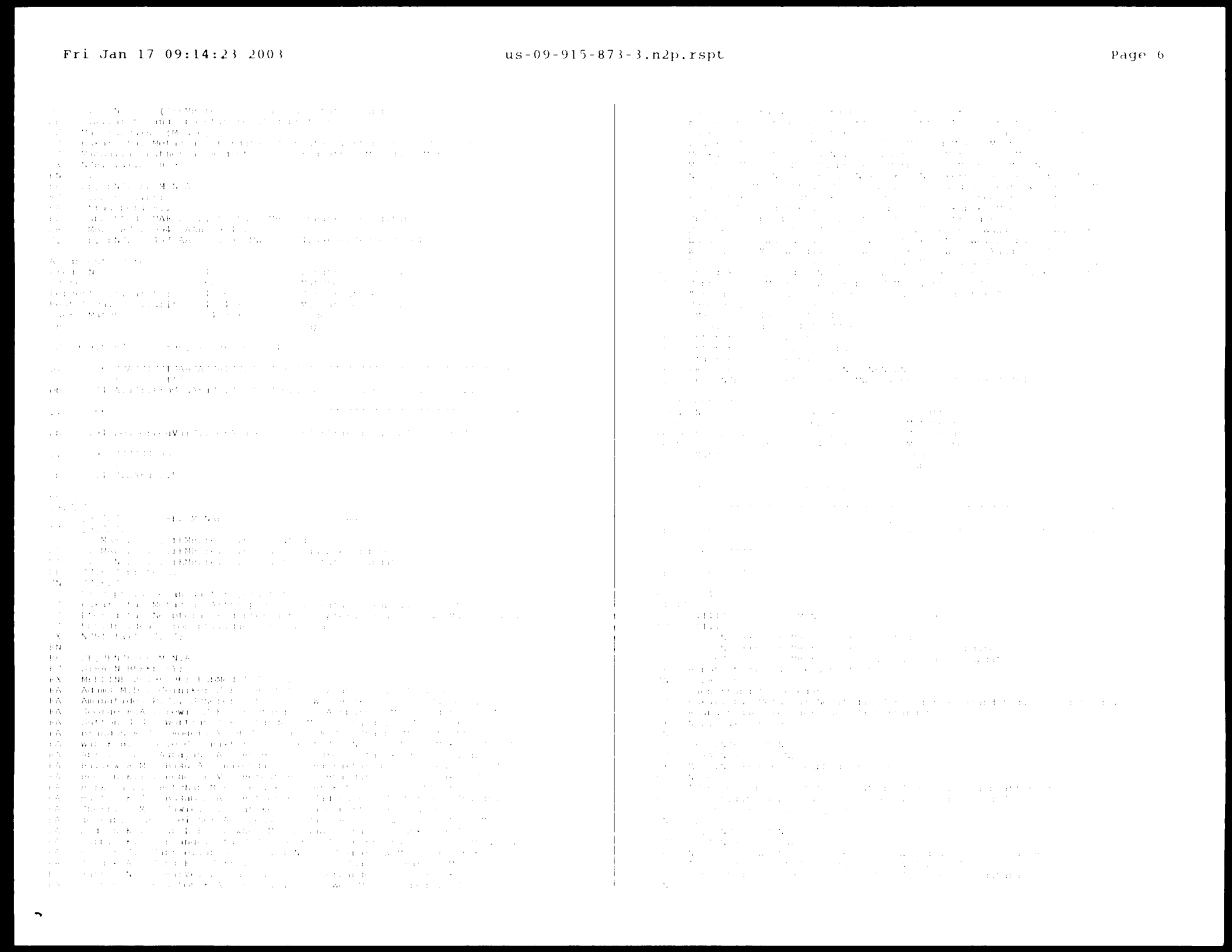
```
Alignment Scores:
Pred. No.:                0.798          Length:          257
Score:                    67.00          Matches:          14
Percent Similarity:       65.22%         Conservative:      1
Best Local Similarity:    60.87%         Mismatches:        6
Query Match:              39.64%         Indels:            2
DB:                       4              Gaps:              1

US-09-915-873-3 (1-96) x Q4H7T3 (1-257)

Qy      79 CGACGACGACCATCACGACGACGACACGCGACGCGAAGTCGTCTTCTTCACTATTCTCTTCATCA 11
           |||   ||||||   ||||||      |||      ||||||   ||||||||:::|||
Db     153 ArgAlaGlyArgArgGlyArg------ArgAlaGlyAlaGlySerSerProSerGlySerAlaGly 17

Qy      10 CCTGCATGG 2
           |||||||
Db     171 AlaAlaTrp 173

RESULT 5
Q9X8A4
ID   Q9X8A4      PRELIMINARY;      PRT;   386 AA.
AC   Q9X8A4;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Hypothetical protein SCO4528.
GN   SCO3528 OR SCE2.09.
OS   Streptomyces coelicolor.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=1902;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RA   Seeger K.J., Harris D.;
RL   Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RA   Bentley S.D., Parkhill J., Barrell B.G., Rajandream M.A.;
RL   Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RX   MEDLINE=97000351; PubMed=8843436;
RA   Redenbach M., Kieser H.M., Denapaite D., Eichner A., Cullum J.,
RA   Kinashi H., Hopwood D.A.;
RT   "A set of ordered cosmids and a detailed genetic and physical map for
RT   the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL   Mol. Microbiol. 21:77-96(1996).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2) / M145;
RA   Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA   Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA   Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA   Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA   Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S.,
RA   Rabbinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA   Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA   Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA   Hopwood D.A.;
RT   "Complete genome sequence of the model actinomycete Streptomyces
RT   coelicolor A3(2).";
RL   Nature 417:141-147(2002).
DR   EMBL; AL049645; CAB11068.1; -.
DR   InterPro; IPR001198; DnaB_helicase.
DR   Pfam; PF00772; DnaB; 2.
KW   Hypothetical protein.
SQ   SEQUENCE   386 AA;  40930 MW;  BB811639AD7AF76A CRC64;

Alignment Scores:
Prod. No.:                0.809          Length:          386
```







```
DR   EMBL; AE003434; AAF46035.1; -.
DR   FlyBase; FBgn0029766; CG15784.
DR   InterPro; IPR002395; Kininogen.
DR   PRINTS; PR00334; KININOGEN.
SQ   SEQUENCE   554 AA;  62429 MW;  9CE2F80A7A[illegible] CRC64;

Alignment Scores:
Pred. No.:              3.03        Length:        [illegible]
Score:                  63.00       Matches:       [illegible]
Percent Similarity:     51.72%      Conservative:  2
Best Local Similarity:  44.83%      Mismatches:    14
Query Match:            35.39%      Indels:        0
DB:                     5           Gaps:          0

US-09-915-873-3 (1-96) x Q9W4C1 (1-554)

Qy       9 CCTCCTCAACACCATCCTCAACAAGCACT[illegible]
             ||||||      :::    |||              |||···      ||||:|:||||
Db     470 GlyProArgPheGlnHisGlnProArgPro[illegible]

Qy      69 CGGGCTGAGCAGCAACCTCACCGCCGG 95
                ||||||       |||    ||||||
Db     490 ProAlaGluPheProProGlyArgArg 498

RESULT 10
Q9P682
ID   Q9P682      PRELIMINARY;      PRT;   803 AA.
AC   Q9P682;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Related to PCF11 component of pre mRNA 3' end processing factor CF
DE   I.
GN   B1D1.390.
OS   Neurospora crassa
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC   Sordariales; Sordariaceae; Neurospora.
OX   NCBI_TaxID=5141;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Schulte U., Aign V., Hoheisel J., Brandt P., Fartmann B., Holland R.,
RA   Nyakatura G., Mewes H.W., Mannhaupt G.;
RL   Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   German Neurospora genome project;
RL   Submitted (NOV-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL355927; CAB91288.2; -.
DR   InterPro; IPR002965; P_rich_extensn.
DR   PRINTS; PR01217; PRICHEXTENSN.
SQ   SEQUENCE   803 AA;  84196 MW;  F99E5AAE[illegible]3BC1E5B CRC64;

Alignment Scores:
Pred. No.:              3.61        Length:        [illegible]
Score:                  62.50       Matches:       15
Percent Similarity:     50.00%      Conservative:  [illegible]
Best Local Similarity:  50.00%      Mismatches:    8
Query Match:            35.11%      Indels:        7
DB:                     3           Gaps:          2

US-09-915-873-3 (1-96) x Q9P682 (1-803)

Qy      12 CCTGAACACGATGGTCAACAAGCACTTCC[illegible]
              |||   |||    |||         |||   ||||||          ||||||||
Db     198 ProProHisGlyGlyAsnThrAlaPheProLeu---- ArgProHisAsnGly[illegible]Asp 215

Qy      63 ------CCTCCTCGGGCTGAGCAGCAACCT 86
                 |||||||||||        ||||||
Db     216 ValLeuProProArgAlaThrThrGlnPro 225

RESULT 11
Q9CU50
ID   Q9CU50      PRELIMINARY;      PRT;   279 AA.
AC   Q9CU50;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   C330049E11RIK protein (Fragment).
GN   C330049E11RIK.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; [illegible]
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; M[illegible]
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=MEDULLA [illegible] NIATA;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., [illegible]
RA   [illegible]
RA   [illegible]
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush [illegible]
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio [illegible]
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., [illegible]
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., W[illegible]
RA   [illegible] Kawaji H., Konagaya A.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
DR   EMBL; AK019113; BAB31076.1; -.
DR   HSSP; P02671; 1FZD.
DR   MGD; MGI:1917970; C330049E11Rik.
DR   InterPro; IPR002181; Fibrinogen_C.
DR   Pfam; PF00147; fibrinogen_C; 1.
DR   SMART; SM00186; FBG; 1.
DR   PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; 1.
FT   NON_TER       1      1
FT   NON_TER     279    279
SQ   SEQUENCE   279 AA;  [illegible] MW;  [illegible] CRC64;

Alignment Scores:
Pred. No.:              4.1         Length:        [illegible]
Score:                  62.00       Matches:       [illegible]7
Percent Similarity:     47.92%      Conservative:  [illegible]
Best Local Similarity:  40.48%      Mismatches:    [illegible]
Query Match:            34.83%      Indels:        [illegible]
DB:                     11          Gaps:          [illegible]

US-09-915-873-3 (1-96) x Q9CU50 (1-279)

Qy       [illegible]
Db       2 AlaGlyProGluArg[illegible]
Qy      [illegible]
Db      22 [illegible]SerLeuValGluSerAlaSerAsnThrSerArgAlaLeuAsp[illegible]
Qy      [illegible]
Db      41 GlnArg 43

RESULT 12
Q8K026
ID   Q8K026      PRELIMINARY;      PRT;   [illegible] AA.
AC   Q8K026;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
```



```
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Waterston R.;
RT   "Direct Submission.";
RL   Submitted (SEP 2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF038605; AAB92020.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   698 AA;  74139 MW;  A6DB9898A4E986EA CRC64;

Alignment Scores:
Pred. No.:              4.98          Length:          698
Score:                  61.50         Matches:         14
Percent Similarity:     41.46%        Conservative:    8
Best Local Similarity:  34.15%        Mismatches:      11
Query Match:            34.55%        Indels:          15
DB:                     5             Gaps:            2

US-09-915-873-3 (1-96) x 044447 (1-698)

Qy       3 CATGCACCC1--------          CCCGAACACGAACCCAACAACACCCCCT  ---  41
           |||   |||                  |||   |||   ||:   |||   :::|||
Db     602 HisProGlyMetAlaGlyMetProProMetHisProGlyMetGlnGlyMetProProMet 642

Qy      42 ------            --CCCCTGTCCCCCCCCA-CCCCCCCCCCCCCCCCCCCCCCC 83
                             ::::||      ||||||   ||.|||::      |||
Db     622 HisProGlyMetGlnGlyMetProGlyMetProHisGlyProProGlnProProGlnAsn 671

Qy      84 CCT 86
           |||
Db     642 Pro 642

RESULT 15
Q95JU7
ID   Q95JU7      PRELIMINARY;      PRT;   728 AA.
AC   Q95JU7;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN 2002 (TrEMBLrel. 21, Last annotation update)
DE   Hypothetical 83.4 kDa protein (Fragment).
OS   Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9541;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=TESTIS;
RA   Hashimoto K., Osada N., Hida M., Kusuda J., Tanuma R., Hirai M.,
RA   Terao K., Sugano S.;
RT   "Isolation of novel full length cDNA clones from macaque testis cDNA
RT   libraries.";
RL   Submitted (AUG 2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB070081; BAB63026.1; -.
DR   InterPro; IPR000315; Znf_Bbox.
DR   Pfam; PF00643; zf-B_box; 1.
KW   Hypothetical protein.
FT   NON_TER     728     728
SQ   SEQUENCE   728 AA;  83411 MW;  5442978852461CFE CRC64;

Alignment Scores:
Pred. No.:              5.87          Length:          728
Score:                  61.00         Matches:         11
Percent Similarity:     60.87%        Conservative:    5
Best Local Similarity:  47.83%        Mismatches:      9
Query Match:            36.09%        Indels:          0
DB:                     6             Gaps:            0

US-09-915-873-3 (1-96) x Q95JU7 (1-728)

Qy      90 CCTCAGGTTGCTGCTCAGCCCGAGAGAGGATGAGCACCTACACAAAAAGAAAGTTT 71
           |||:::|||                 |||:::|||||||||   :::|||||||.|
Db     320 GlyAspValGlnCysSerGlnAsnGlnAsnAspGlnAspSerGlnGlyGlnGlyHisLys 379

Qy      30 CTTCAACAC 22
           .|:  .||
Db     380 Val Thr His 388

Search completed: January 17, 2003, 09:11:19
Job time : 84 secs
```

GenCore version 5.1.3

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 17, 2003, 07:01:12 ; Search time 10 Seconds
(without alignments)
564.920 Million amino acid updates/sec

Title: US-09-915-873-3
Perfect score: 178
Sequence: 1 accatgcaggtcctgaacac.....gcagcaacctcacccgcagc 96

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 525148

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV xlh
-Q=/cgn2_1/tmp/tmp_app_query/1461203_1_73313_1.fasta 1,96
-DB=Issued_Patents_AA -OFMT fastam -SUFFIX=n2p.rai -MINMATCH=0.1 -1=3*3
-LOOPEXT=0 -UNITS=bits -START 1 -END=-1 -MATRIX blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 ALIGN 15
-MODE=LOCAL -OUTFMT pto -NORM=ext -HEAPSIZE 500 MINLEN 0 MAXLEN=2000000000
-USER=US09915873_@CGN_1_1_3_@runat_14012003_171644_13340 -NCPU 6 -ICPU=3
-NO_XLPXY -NO_MMAP -TARGETQUERY -NPT_SCOPES 0 -WAIT -LONG=3 -DEV_TIMEOUT 120
-WARN_TIMEOUT=30 -THREADS=1 -XGAP=10.0 -XGAPEXT=0.5 -YGAP=10.0 -YGAPEXT=0.5
-YGAPOP=10.0 -YGAPEXT=0.5 -DELOP=6.0 -DELEXT=7.0

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/PA_COMB.pep *
2: /cgn2_6/ptodata/1/iaa/PB_COMB.pep *
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep *
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep *
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep *
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep*

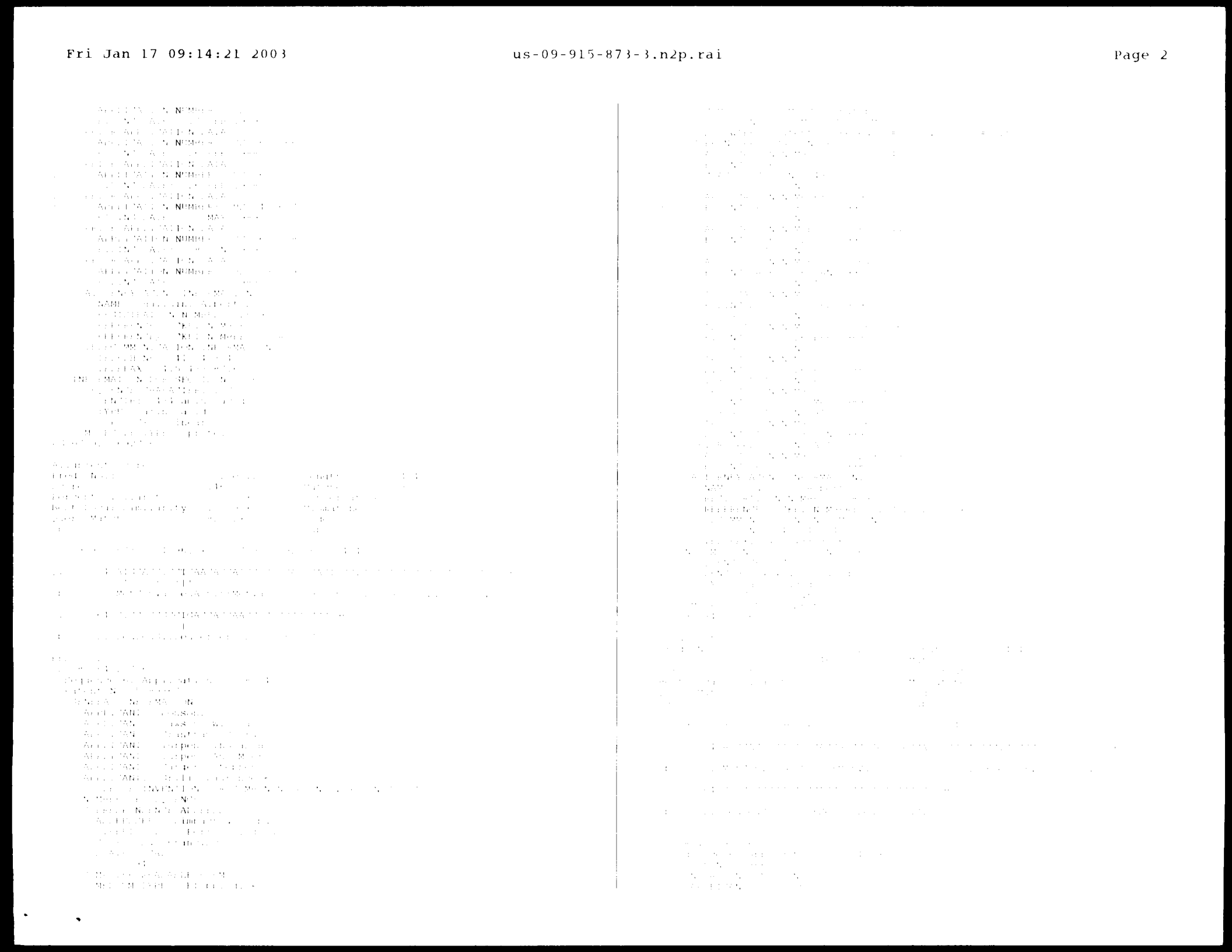
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	146	82.0	434	1	US-07-922-492A-6	Sequence 6, Appl
2	146	82.0	434	1	US-08-08-194-277-4	Sequence 4, Appl
3	146	82.0	434	2	US-09-09-122-922-4	Sequence 4, Appl
4	146	82.0	434	3	US-08-484-441-6	Sequence 6, Appl
5	146	82.0	434	4	US-08-483-502-6	Sequence 6, Appli
6	146	82.0	434	1	US-09-226-651A-6	Sequence 6, Appl
7	117	65.7	25	3	US-09-023-339-5	Sequence 5, Appl
8	117	65.7	428	1	US-07-973-424A-4	Sequence 4, Appli
9	117	65.7	428	1	US-08-343-386-4	Sequence 4, Appl
10	117	65.7	428	4	US-09-072-435-4	Sequence 4, Appl
11	117	65.7	428	4	US-09-072-917A-4	Sequence 4, Appl
12	71.5	40.2	25	6	5198415-24	Patent No. 5198
13	71.5	40.2	43	1	US-07-922-492A-12	Sequence 12, Appl
14	71.5	40.2	43	1	US-08-08-194-277-12	Sequence 12, Appl
15	71.5	40.2	43	1	US-08-08-194-277-12	Sequence 12, Appl
16	71.5	40.2	43	1	US-08-08-194-277-12	Sequence 12, Appl
17	71.5	40.2	43	1	US-08-08-194-277-12	Sequence 12, Appl
18	71.5	40.2	43	1	US-08-08-194-277-12	Sequence 12, Appl
19	71.5	40.2	43	3	US-08-484-441-12	Sequence 12, Appl
20	71.5	40.2	43	3	US-08-484-441-12	Sequence 12, Appl
21	71.5	40.2	1001	4	US-08-483-502-27	Sequence 27, Appl
22	71.5	40.2	1001	4	US-08-483-502-27	Sequence 27, Appl
23	71.5	40.2	1001	4	US-08-483-502-27	Sequence 27, Appl
24	71.5	40.2	1000	4	US-08-483-502-27	Sequence 27, Appl
25	69	38.8	147	3	US-08-484-441-2	Sequence 2, Appl
26	69	38.8	1417	1	US-08-08-194-277-4	Sequence 4, Appl
27	68	38.2	2084	1	US-08-08-194-277-4	Sequence 4, Appl
28	68	38.2	2084	1	US-08-08-194-277-4	Sequence 4, Appl
29	68	38.2	2084	2	US-09-09-122-922-4	Sequence 4, Appl
30	68	38.2	2084	3	PCT-US94-07271-4	Sequence 4, Appl
31	68	38.2	4278	3	US-08-484-441-1	Sequence 1, Appl
32	68	38.2	4278	4	US-08-483-502-1	Sequence 1, Appl
33	67	37.6	1418	1	US-08-08-194-277-4	Sequence 4, Appl
34	67	37.6	1418	3	US-08-484-441-4	Sequence 4, Appl
35	67	37.6	2084	3	US-08-484-441-5	Sequence 5, Appl
36	66.5	37.4	1001	1	US-08-08-194-277-27	Sequence 27, Appl
37	66	37.1	1418	1	US-08-08-194-277-4	Sequence 4, Appl
38	66	37.1	1418	4	US-08-483-502-4	Sequence 4, Appl
39	66	37.1	1418	4	US-08-483-502-4A-4	Sequence 4, Appl
40	66	37.1	1418	4	US-08-483-502-4	Sequence 4, Appl
41	66	37.1	1418	4	US-09-072-786-4	Sequence 4, Appl
42	65.5	36.8	1417	1	US-08-08-194-277-4	Sequence 4, Appl
43	65	36.5	1418	3	US-08-484-441-4	Sequence 4, Appl
44	65	36.5	1418	4	US-08-483-502-4A-4	Sequence 4, Appl
45	65	36.5	1418	4	US-08-483-502-4A-4	Sequence 4, Appl

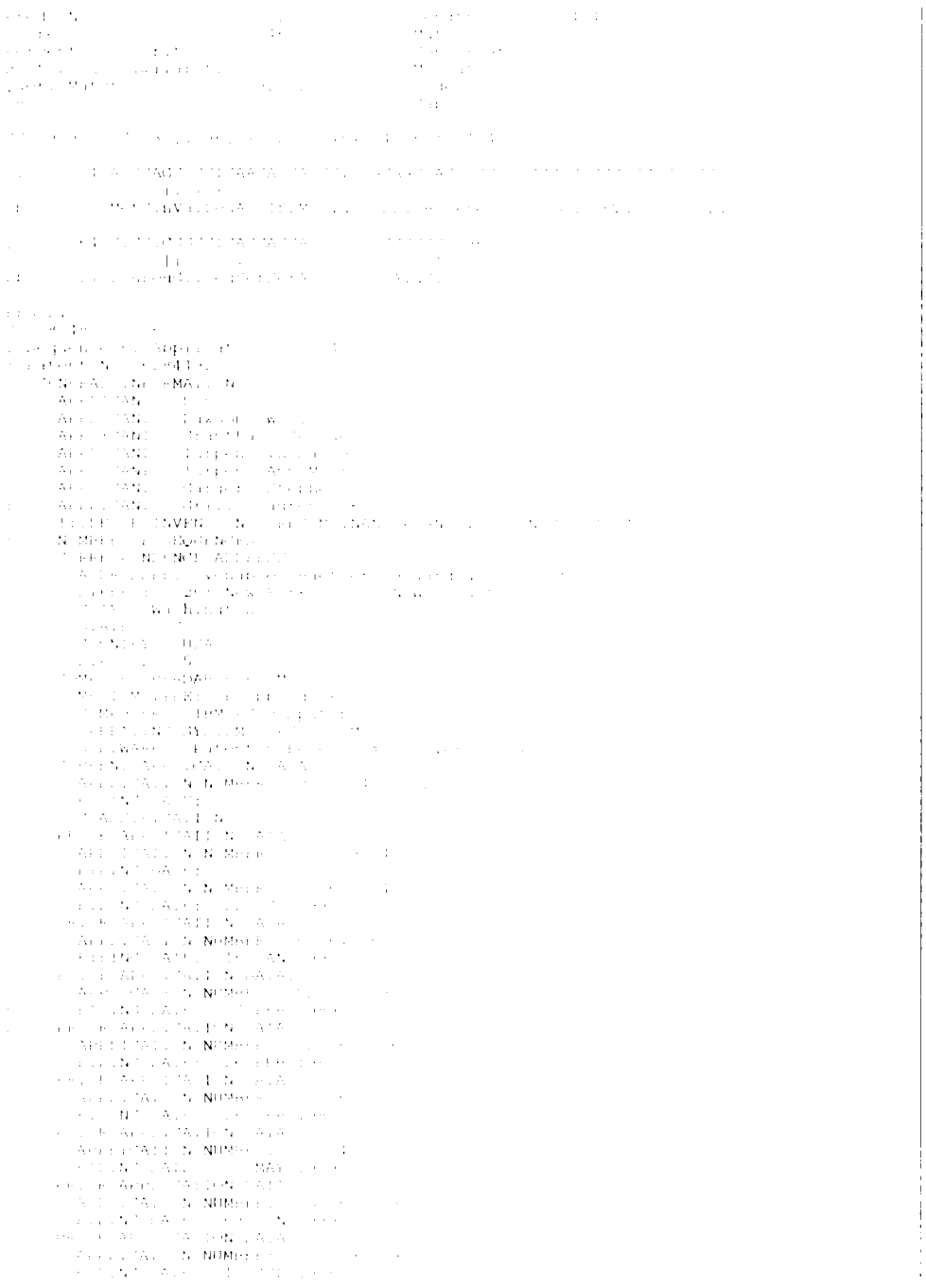
ALIGNMENTS

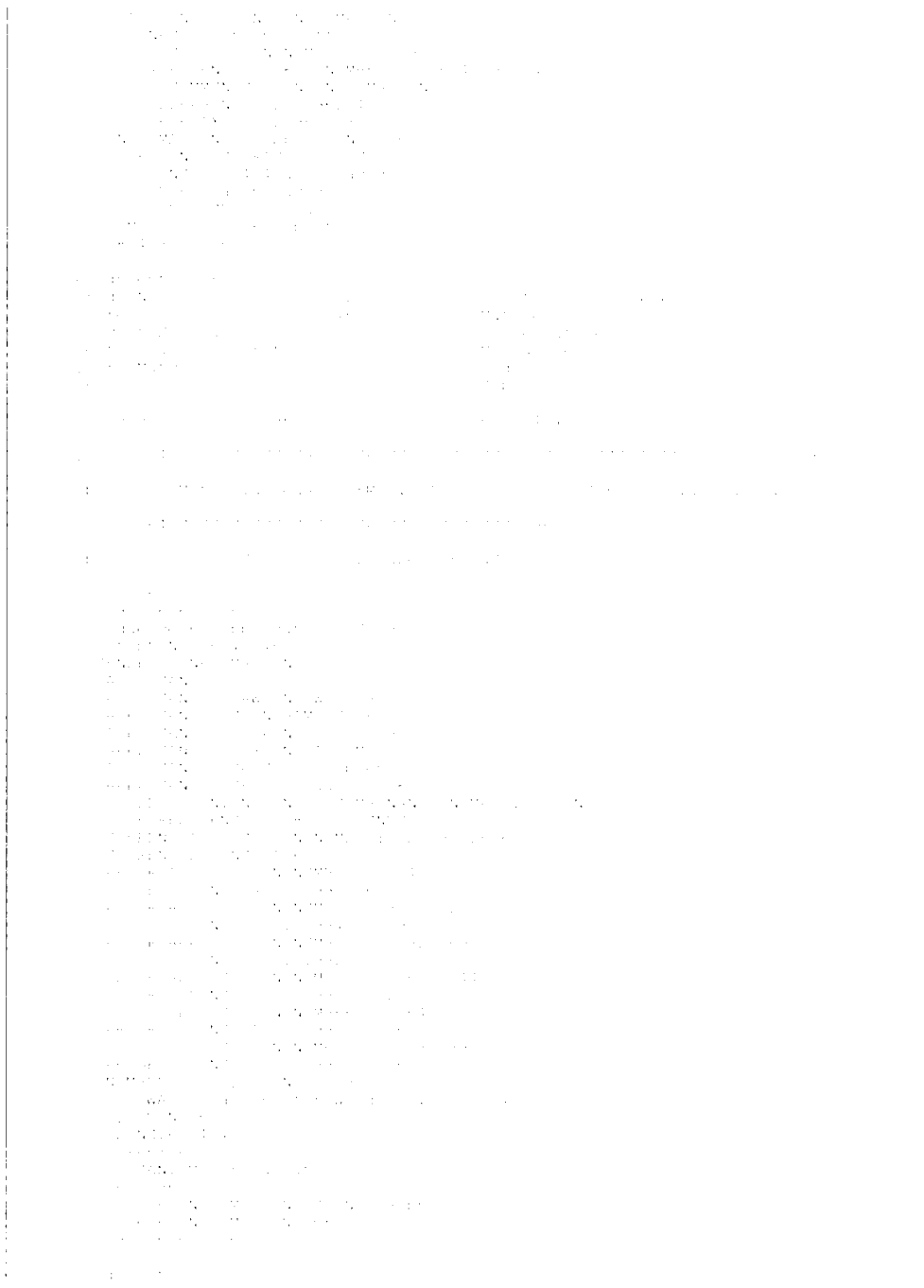
RESULT 1
US-07-922-492A-6
; Sequence 6, Application US/07922492A
; Patent No. 5816251
; GENERAL INFORMATION:
; APPLICANT: Ronson, Jon
; APPLICANT: Dawson, Willie J.
; APPLICANT: Bradham, Peter A.
; APPLICANT: Tarpley, Thomas M.
; APPLICANT: Tarpley, Ann Myers
; APPLICANT: Tarpley, Stephen A.
; APPLICANT: Grieb, Lawrence K.
; TITLE OF INVENTION: Novel Compounds, Compositions and Methods for Tumor Therapy
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Limbach & Limbach
; STREET: 2001 Ferry Building
; CITY: San Francisco
; STATE: CA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: FastSEQ for Windows Version 1.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/922,492
; FILING DATE: 31-JUL-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/
; FILING DATE: 22-JUL-
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/
; FILING DATE: 16-JAN-
; PRIOR APPLICATION DATA:



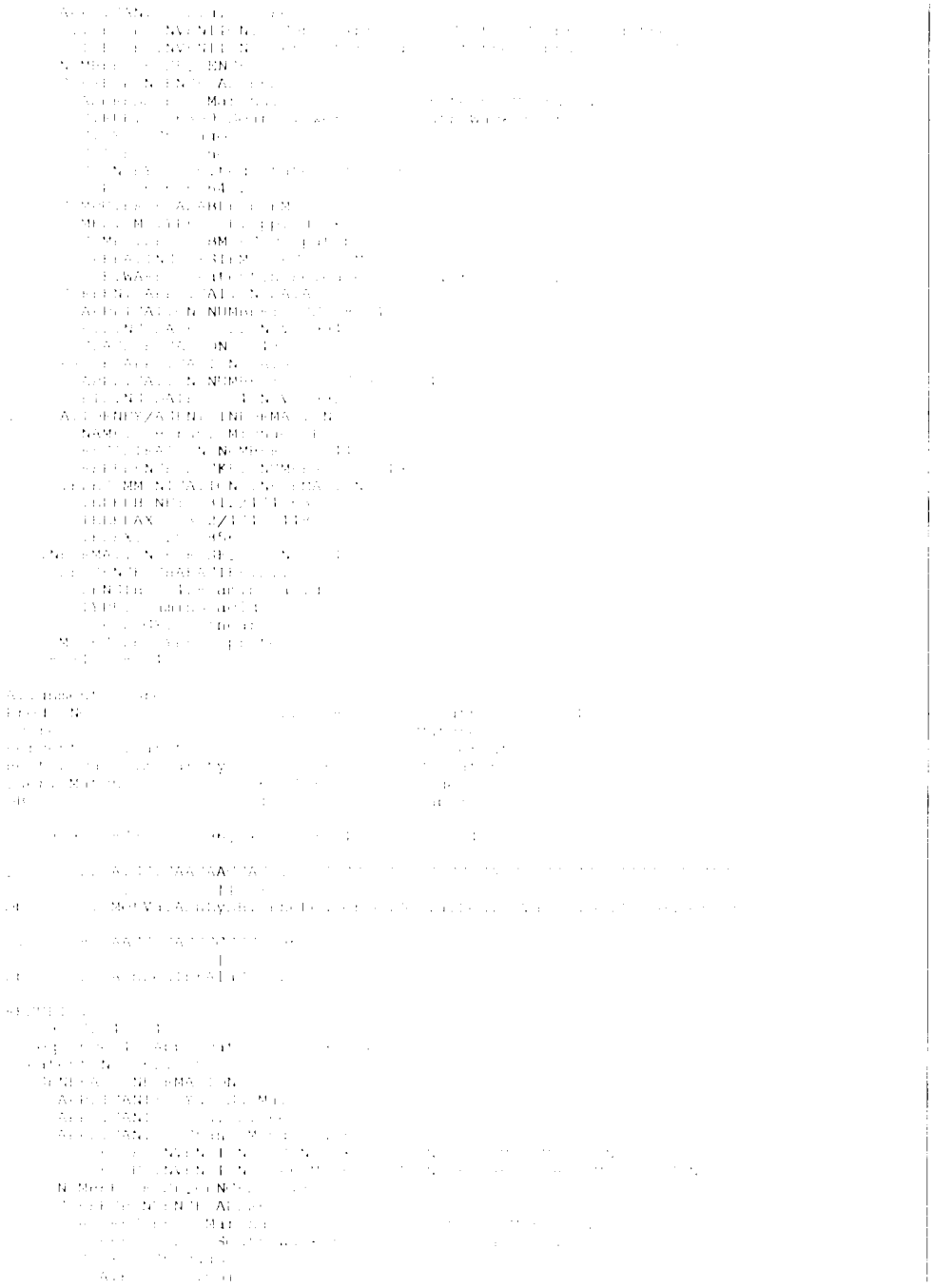
```
;     APPLICANT:  Dawson, William O.
;     APPLICANT:  Grantham, George L.
;     APPLICANT:  Turpen, Thomas H.
;     APPLICANT:  Turpen, Ann Myers
;     APPLICANT:  Garger, Stephen J.
;     APPLICANT:  Grill, Laurence K.
;     TITLE OF INVENTION:  RECOMBINANT PLANT VIRAL NUCLEIC ACIDS
;     NUMBER OF SEQUENCES:  19
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Fennie & Edmonds
;       STREET:  1155 Avenue of the Americas
;       CITY:  New York
;       STATE:  New York
;       ZIP:  10036
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC DOS/MS-DOS
;       SOFTWARE:  Patent in Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/492,920
;       FILING DATE:  07-JUNE-1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 184,237
;       FILING DATE:  19-JAN-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 600,244
;       FILING DATE:  22-OCT-1990
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 641,617
;       FILING DATE:  16-JAN-1991
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 310,881
;       FILING DATE:  17-FEB-1989
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 160,766
;       FILING DATE:  26-FEB-1988
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 160,771
;       FILING DATE:  26-FEB-1988
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 347,637
;       FILING DATE:  05-MAY-1989
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 363,138
;       FILING DATE:  08-JUN-1989
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 219,279
;       FILING DATE:  15-JUL-1988
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Halluin, Albert P.
;       REGISTRATION NUMBER:  28,957
;       REFERENCE/DOCKET NUMBER:  8129 112
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  415-854-3660
;       TELEFAX:  415-854-3694
;   INFORMATION FOR SEQ ID NO:  6:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  434 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-482-920-6

Alignment Scores:
Pred. No.:                3.2le-12        Length:          434
Score:                    146.00          Matches:         31
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              82.02%          Indels:          0
DB:                       2               Gaps:            0

US-09-915-873-3 (1-96) x US-08-482-920-6 (1-434)
```

```
Qy      4 APGAGSETPGAAVALQKEQQAALSTVPEVATQVVSQVLIVEQAIEADIYGKVTLATLTLLL
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MSTGDVLISAAHTHSVLAAMLGAEGIAFTGTSLGSEIAVMRSRVWYDQTEMAVR

Qy     64 LQQLQQQMELVLNLRLRELEMSVDRLSSYVDFAFVDFQEVRNTLGGNIPTDAPSRIKFVD
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     21 LQQLQQQMELVLNLRLRELEMSVDRLSSYVDFAFVDFQEVRNTLGGNIPTDAPSRIKFVD

RESULT 4
US-08-482-920-6
; Sequence 6, Application US/08482920
;   GENERAL INFORMATION:
;       APPLICANT: Dawson, Wil
;                  Dawson, Wil
;                  Grantham, G
;                  Turpen, Tho
;                  Turpen, Ann
;                  Garger, Ste
;                  Grill, Lau
;       TITLE OF INVENTION: RECOMBINANT PLANT VIRAL NUCLEIC ACI
;       NUMBER OF SEQUENCES: 19
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Fen
;            STREET: 1155
;            CITY: New Yo
;            STATE: NY
;            ZIP: 10036
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Flo
;            COMPUTER: IBM P
;            OPERATING SYSTE
;            SOFTWARE: Pate
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/08/482,920
;            FILING DATE: 07-JUN-1995
;            CLASSIFICATION:
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER:
;            FILING DATE: Unknown
;            APPLICATION NUMBER:
;            FILING DATE:
;            APPLICATION NUMBER:
;            FILING DATE:
;            APPLICATION NUMBER:
;            FILING DATE:
;            APPLICATION NUMBER:
;            FILING DATE:
;            APPLICATION NUMBER:
;            FILING DATE:
;            APPLICATION NUMBER:
;            FILING DATE:
;            APPLICATION NUMBER:
;            FILING DATE:
;       ATTORNEY/AGENT INFORMATION:
;            NAME:
;            REGISTRATION NUMBER:
;            REFERENCE/DOCKET NUMBER:
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE:
;            TELEFAX:
;   INFORMATION FOR SEQ ID NO:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH:
;            TYPE:
;            TOPOLOGY:
;       MOLECULE TYPE:
;       SEQUENCE DESCRIPTION:
US-08-484-841-6

Alignment Scores:
```

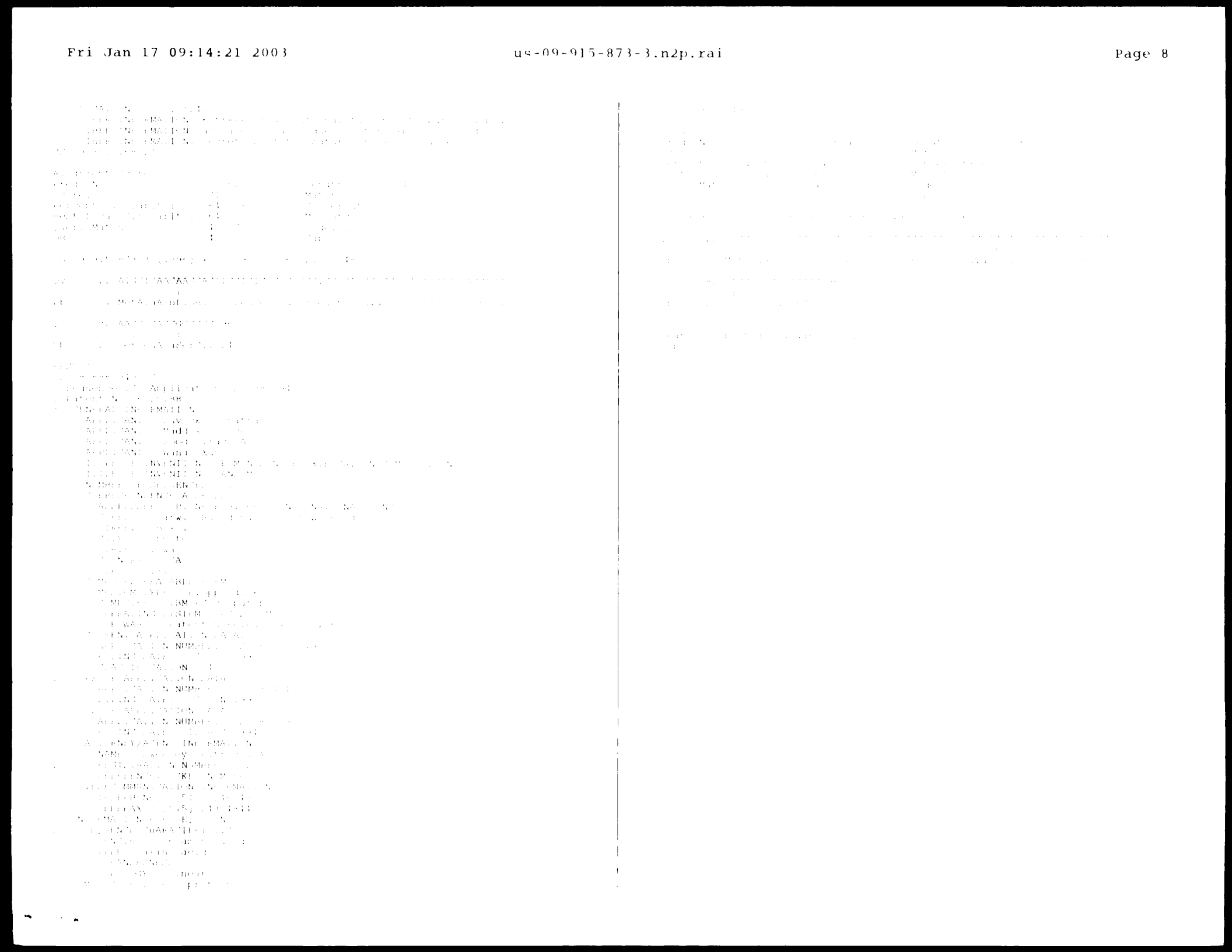




```
Pred. No.:              3.21e-12        Length:         434
Score:                  146.00          Matches:        31
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            82.02%          Indels:         0
DB:                     4               Gaps:           0

US-09-915-873-3 (1-96) x US-09-726-651A-6 (1-434)

Qy       4 ATGCAGGTCCTGAACACGATGGTCAACAAGCACTTCCTCTTCGGTGCCGTGAACGAGCATCAGTTC 69
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MetGlnValLeuAsnThrMetValAsnLysHisPheLeuSerLeuSerValLeuIleVal 60

Qy      64 CTCCTCGGGGCTGAGCAGCAACCTCACGGCGGGC 96
           |||||||||||||||||||||||||||||||||
Db      21 LeuLeuGlyLeuSerSerAsnLeuThrAlaGly 53

RESULT 7
US-09-023-339-5
; Sequence 5, Application US/09023339
; Patent No. 6127145
;  GENERAL INFORMATION:
;    APPLICANT:  Sutliff, Thomas D.
;    APPLICANT:  Rodriquez, Raymond L.
;    TITLE OF INVENTION:  Production Of Alpha-1-Antitrypsin
;    TITLE OF INVENTION:  In Plants
;    NUMBER OF SEQUENCES:  22
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dehlinger & Associates
;      STREET:  P.O. Box 60850
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94306
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/023,339
;      FILING DATE:  13-FEB-1998
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/037,991
;      FILING DATE:  13-FEB-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Petithory, Joanne R
;      REGISTRATION NUMBER:  P42,995
;      REFERENCE/DOCKET NUMBER:  9665.0001.00
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-324-0880
;      TELEFAX:  650-324-0960
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  25 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    IMMEDIATE SOURCE:
;      CLONE:  RAmylA signal peptide
US-09-023-339-5

Alignment Scores:
Pred. No.:              1.92e-08        Length:         25
Score:                  117.00          Matches:        25
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            65.73%          Indels:         0
DB:                     3               Gaps:           0

US-09-915-873-3 (1-96) x US-09-023-339-5 (1-25)

Qy      22 ATGGTCAACAAGCACTTCCTCTTCGGTGCCGTGAACGAGCATCAG 67
           |||||||||||||||||||||||||||||||||||||||||||||
Db       1 MetValAsnLysHisPheLeuSerLeuSerValLeuIleVal 40

Qy      68 AACTTCACCGCCGGC 82
           |||||||||||||||
Db      21 AsnLeuThrAlaGly 37

RESULT 8
US-09-973-324A-4
; Sequence 4, Application US/09973324A
; Patent No. 6466...
;  GENERAL INFORMATION:
;    APPLICANT:  Yu, Su-Mei
;    APPLICANT:  Lin, Li-Pei
;    TITLE OF INVENTION:  Gene Expression System Comprising ...
;    TITLE OF INVENTION:  Inducible Seed-Specific Plant Promoter ...
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Patentin Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/973,324A
;      FILING DATE:  04-NOV-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Borun, Michael ...
;      REGISTRATION NUMBER:  ...,...
;      REFERENCE/DOCKET NUMBER:  31 PA
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-...
;      TELEFAX:  312/474-...
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4.. amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-973-324A-4

Alignment Scores:
Pred. No.:              ...e-08         Length:         ...
Score:                  ...             Matches:        ...
Percent Similarity:     ...             Conservative:   ...
Best Local Similarity:  ...             Mismatches:     ...
Query Match:            65.73%          Indels:         0
DB:                     ...             Gaps:           ...

US-09-915-873-3 (1-96) x US-09-973-324A-4 (1-...

Qy      22 ATGGTCAACAAGCACTTCCTCTTCGGTGCCGTGAACGAGCATCAGTTCACCGCCGGC 81
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       . MetValAsnLysHisPheLeuSerLeuSerValLeuIleValLeuLeuGlyLeuSer ...

Qy      82 AACTTCACCGCCGGC 96
           |||||||||||||||
Db      ... AsnLeuThrAlaGly ...

RESULT 9
US-08-...-...A-4
; Sequence 4, Application US/...
; Patent No. 5...
;  GENERAL INFORMATION:
;    APPLICANT:  Yu, Su-Mei
```



```
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/072,917A
;       FILING DATE:
;       CLASSIFICATION:  800
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/509,962
;       FILING DATE:  01-AUG-1995
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Gass, David A.
;       REGISTRATION NUMBER:  38,153
;       REFERENCE/DOCKET NUMBER:  29123/31257
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (312) 474-6300
;       TELEFAX:  (312) 474-0448
;   INFORMATION FOR SEQ ID NO:  4:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  428 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-09-072-917A-4

Alignment Scores:
Pred. No.:                2.22e-09        Length:          428
Score:                    117.00          Matches:         25
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              65.73%          Indels:          0
DB:                       4               Gaps:            0

US-09-915-873-3 (1-96) x US-09-072-917A-4 (1-428)

Qy      22 ATGGTCAACAAGCACTTCCTCTCCGTCCGTCCGTCTCATCGTCCTCCTCCGGCCTCTACCAC 81
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MetValAsnLysHisPheLeuSerLeuSerValLeuIleValLeuLeuGlyLeuSerSer 20

Qy      82 AACCTCACCGCCGGC 96
           ||||||||||||||||
Db      21 AsnLeuThrAlaGly 25

RESULT 12
5198345 24
;Patent No. 5198345
;     APPLICANT: GWYNNE, DAVID I.;BUXTON, FRANCIS P.;PUGHEIL, MARK H.
;DAVIES, ROGER W.;SCAZZOCCHIO, CLAUDIO
;     TITLE OF INVENTION: VECTORS IN USE IN FILAMENTOUS FUNGI
;     NUMBER OF SEQUENCES: 28
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/07/811,404
;       FILING DATE: 20-DEC-1985
;SEQ ID NO:24:
;       LENGTH: 25
5198345-24

Alignment Scores:
Pred. No.:                0.0366          Length:          25
Score:                    71.50           Matches:         16
Percent Similarity:       84.00%          Conservative:    5
Best Local Similarity:    64.00%          Mismatches:      3
Query Match:              40.17%          Indels:          1
DB:                       6               Gaps:            1

US-09-915-873-3 (1-96) x 5198345-24 (1-25)

Qy      22 ATGGTCAACAAGCACTTCCTCTCCGTCCGTCCGTCTCATCGTCCTCCTCCGGCCTCTACCAC 81
           |||   |||||||||   |||||||||||||:::   :::|:|:|:|||||| |||:::
Db       1 MetAlaAsnLysHis---LeuSerLeuSerLeuPheLeuValLeuLeuGlyLeuSerAla 17

Qy      82 AACCTCACCGCCGGC 96
           :::|||   :::|||
Db      18 SerLeuAlaSerGly 24

RESULT 13
US-09-452-159-21
; Sequence 21, Application US/09452159
; Patent No. 6211434
; GENERAL INFORMATION:
;  APPLICANT: Durzick, Jonathan L.
;  APPLICANT: Gillam, Jack E.
;  APPLICANT: Maddox, Joyce E.
;  TITLE OF INVENTION: Amino Acid, Acids Associated
;  TITLE OF INVENTION: Polynucleotides and Related Polypeptides and Methods
;  FILE REFERENCE: 1154
;  CURRENT APPLICATION NUMBER: US/09/452,159A
;  CURRENT FILING DATE: 1999-12-01
;  EARLIER APPLICATION NUMBER: 60/111,000
;  EARLIER FILING DATE: 1998-12-04
;  EARLIER APPLICATION NUMBER: 60/111,000
;  EARLIER FILING DATE: 1998-12-04
;  NUMBER OF SEQ ID NOS: 34
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 21
;   LENGTH: 487
;   TYPE: PRT
;   ORGANISM: Unknown
;   FEATURE:
;   NAME/KEY: SIGNAL
;   LOCATION: (1)...(43)
US-09-452-159-21

Alignment Scores:
Pred. No.:                [illegible]     Length:          [illegible]
Score:                    71.5            Matches:         [illegible]
Percent Similarity:       84.00%          Conservative:    [illegible]
Best Local Similarity:    64.00%          Mismatches:      [illegible]
Query Match:              40.17%          Indels:          [illegible]
DB:                       [illegible]     Gaps:            [illegible]

US-09-915-873-3 (1-96) x [illegible] (1-487)

Qy      22 ATGGTCAACAAGCACTTCCTCTCCGTCCGTCCGTCTCATCGTCCTCCTCCGGCCTCTACCAC 81
           |||   |||||||||   |||||||||||||:::   :::|:|:|:|||||| |||:::
Db       1 MetAlaAsnLysHis   [illegible]

Qy      82 AACCTCACCGCCGGC 96
           ||||||   ||||||
Db      20 SerLeuAlaSerGly 34

RESULT 14
US-09-452-159-21
; Sequence 21, Application US/09452,159A
; Patent No. 6211434
; GENERAL INFORMATION:
;  APPLICANT: Durzick, Jonathan L.
;  APPLICANT: Durzick, Jonathan L.
;  APPLICANT: Eckert, S.
;  APPLICANT: Gillam, Jack E.
;  APPLICANT: Maddox, Joyce E.
;  TITLE OF INVENTION: Amino Acid, Acids Associated
;  TITLE OF INVENTION: Polynucleotides and Related Polypeptides and Methods
;  FILE REFERENCE: 1154
;  CURRENT APPLICATION NUMBER: US/09/452,159A
;  CURRENT FILING DATE: 1999-12-01
;  EARLIER APPLICATION NUMBER: 60/111,000
;  EARLIER FILING DATE: 1998-12-04
;  NUMBER OF SEQ ID NOS: 34
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 21
;   LENGTH: 487
;   TYPE: PRT
;   ORGANISM: Unknown
;   FEATURE:
;   NAME/KEY: SIGNAL
```



GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 17, 2003, 07:01:47 ; Search time 8 Seconds
(without alignments)
477.084 Million cell updates/sec

Title: US-09-915-873-3
Perfect score: 178
Sequence: 1 accatgcaagtcctcaacaa.....acaagcaactcctgaccacaaga 90

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 120991 seqs, 19878514 residues

Total number of hits satisfying chosen parameters: 241982

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=x11
-Q=/cgn2_1/tmp/US09915873_3.rnat_14012003_171645_13495/app_query.fasta_1.2.4
-DB=Published_Applications_AA -QFMT=fastan -SUFFIX=n2p.rapb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-IRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09915873_3_CGN_1_1_2_rnat_14012003_171645_13495
-NCPU=6 -ICPU=3 -NO_XLPXY -NO_MMAP -LARGEQUERY -NPG_SCORES=0 -WAIT -UNGCL=2
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubaa/US09_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubaa/US07_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubaa/PCTUS_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubaa/US06_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubaa/US05_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubaa/US04_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubaa/US04_NEW_PUB.pep:*
10: /cgn2_6/ptodata/2/pubaa/US03_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubaa/US02_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubaa/US01_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubaa/US00_NEW_PUB.pep:*
14: /cgn2_6/ptodata/2/pubaa/US99_PUBCOMB.pep:*

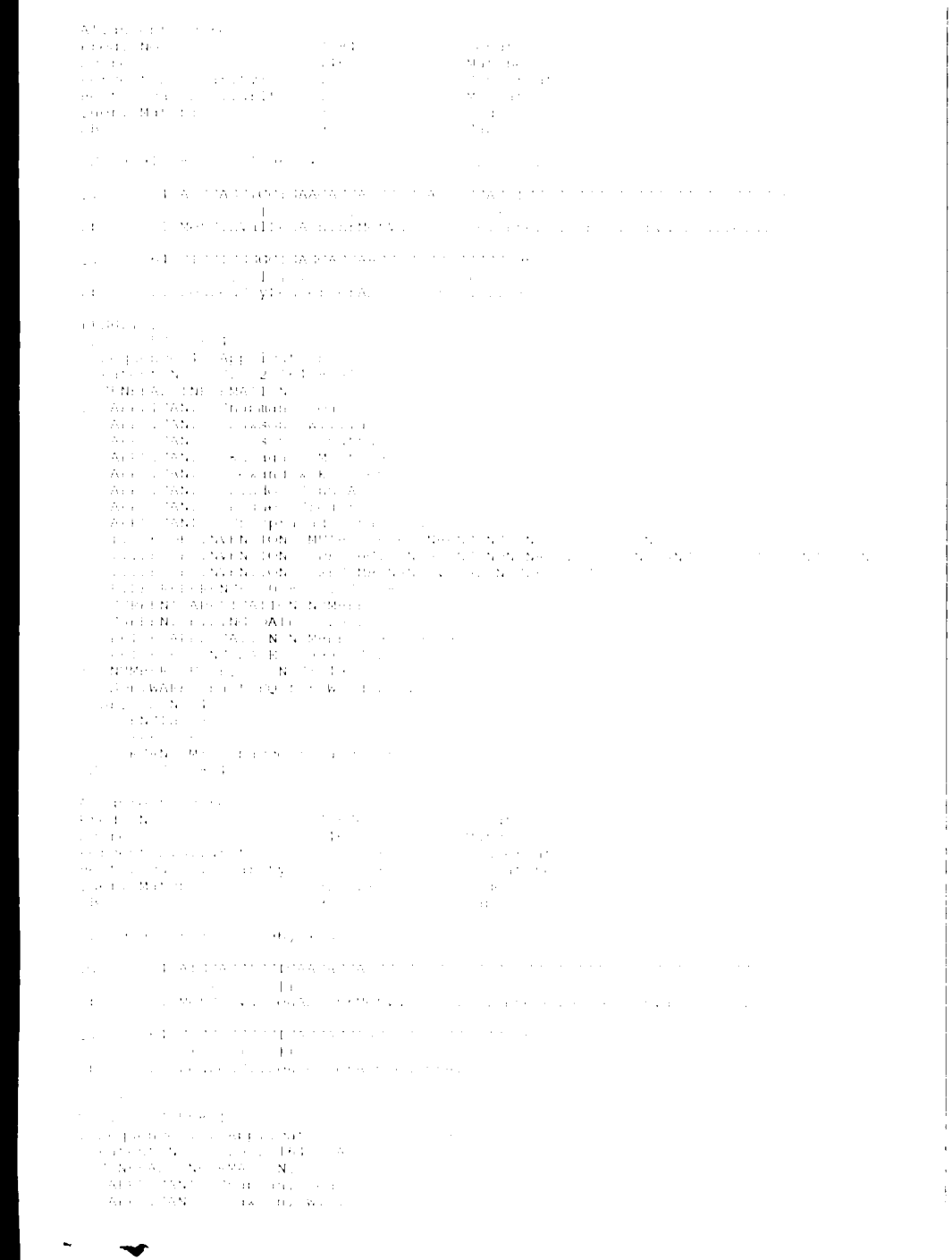
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

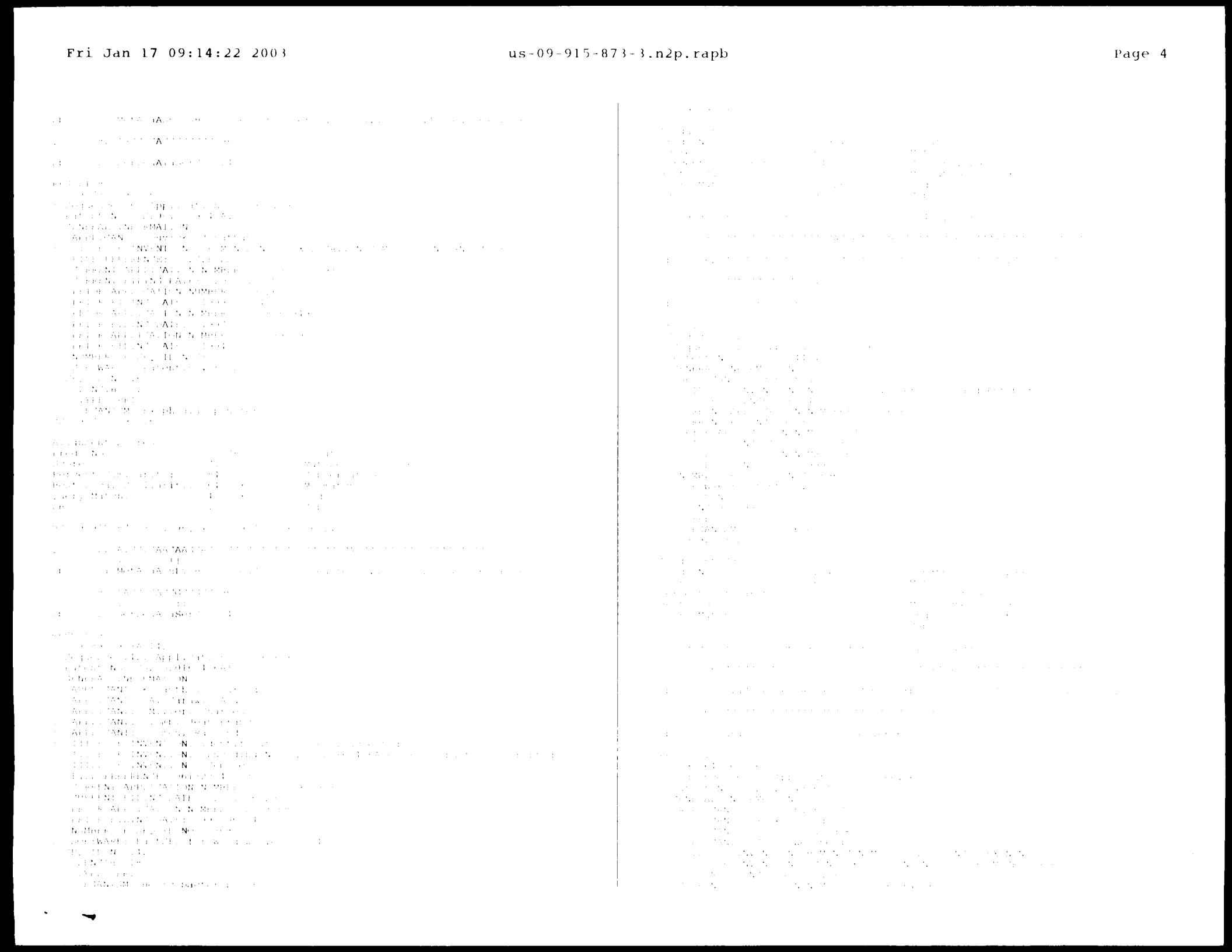
Result No.	Score	% Query Match	Length	DB	ID	Description
1	146	82.0	32	9	US-10-057-558-12	Sequence 12, Appl
2	146	82.0	33	9	US-10-057-558-4	Sequence 4, Appl
3	146	82.0	40	9	US-10-057-558-10	Sequence 10, Appl
4	131.5	73.9	60	9	US-10-057-558-8	Sequence 8, Appl
5	125	70.2	30	9	US-10-057-558-2	Sequence 2, Appl
6	124	69.7	29	9	US-10-057-558-3	Sequence 3, Appl
7	71.5	40.2	30	10	US-10-057-558-17	Sequence 17, Appl
8	71.5	40.2	520	11	US-10-057-558-6	Sequence 6, Appl
9	69	38.8	48	9	US-10-057-558-11	Sequence 11, Appl
10	68	38.2	209	9	US-10-057-558-5	Sequence 5, Appl
11	68	38.2	248	9	US-10-057-558-9	Sequence 9, Appl
12	68	38.2	478	12	US-09-915-873-4	Sequence 4, Appl
13	68	38.2	1000	9	US-10-057-558-7	Sequence 7, Appl
14	67.5	37.9	90	9	US-10-057-558-13	Sequence 13, Appl
15	67.5	37.9	90	9	US-10-057-558-14	Sequence 14, Appl
16	67	37.6	45	9	US-10-057-558-1	Sequence 1, Appl
17	66.5	37.4	60	10	US-10-057-558-16	Sequence 16, Appl
18	66	37.1	33	12	US-09-915-873-3	Sequence 3, Appl
19	66	37.1	45	9	US-10-057-558-15	Sequence 15, Appl
20	66	37.1	51	9	US-10-057-558-18	Sequence 18, Appl
21	66	37.1	60	9	US-10-057-558-19	Sequence 19, Appl
22	66	37.1	120	9	US-10-057-558-20	Sequence 20, Appl
23	66	37.1	1300	9	US-10-057-558-21	Sequence 21, Appl
24	66	37.1	1500	9	US-10-057-558-22	Sequence 22, Appl
25	66	37.1	1500	9	US-10-057-558-23	Sequence 23, Appl
26	66	37.1	1500	9	US-10-057-558-24	Sequence 24, Appl
27	64.5	36.2	2221	9	US-10-057-558-25	Sequence 25, Appl
28	64	36.0	60	10	US-09-915-873-2	Sequence 2, Appl
29	64	36.0	63	11	US-09-915-873-5	Sequence 5, Appl
30	64	36.0	64	11	US-09-915-873-6	Sequence 6, Appl
31	63	35.4	1100	9	US-10-057-558-26	Sequence 26, Appl
32	63	35.4	5000	10	US-09-915-873-1	Sequence 1, Appl
33	63	35.4	5000	10	US-09-915-873-7	Sequence 7, Appl
34	63	35.4	7000	9	US-10-057-558-27	Sequence 27, Appl
35	63	35.4	9000	9	US-10-057-558-28	Sequence 28, Appl
36	62.5	35.1	500	9	US-10-057-558-29	Sequence 29, Appl
37	62	34.8	52	9	US-09-760-842-1587	Sequence 1587, Ap
38	62	34.8	90	9	US-09-760-820-262	Sequence 262, Appl
39	62	34.8	126	9	US-10-057-558-30	Sequence 30, Appl
40	62	34.8	144	9	US-10-057-558-31	Sequence 31, Appl
41	62	34.8	150	9	US-10-057-558-32	Sequence 32, Appl
42	62	34.8	197	10	US-09-760-842-126	Sequence 126, Appl
43	62	34.8	435	10	US-09-760-842-752	Sequence 752, Appl
44	62	34.8	50	9	US-10-057-558-33	Sequence 33, Appl
45	62	34.8	60	10	US-09-915-873-8	Sequence 8, Appl

ALIGNMENTS

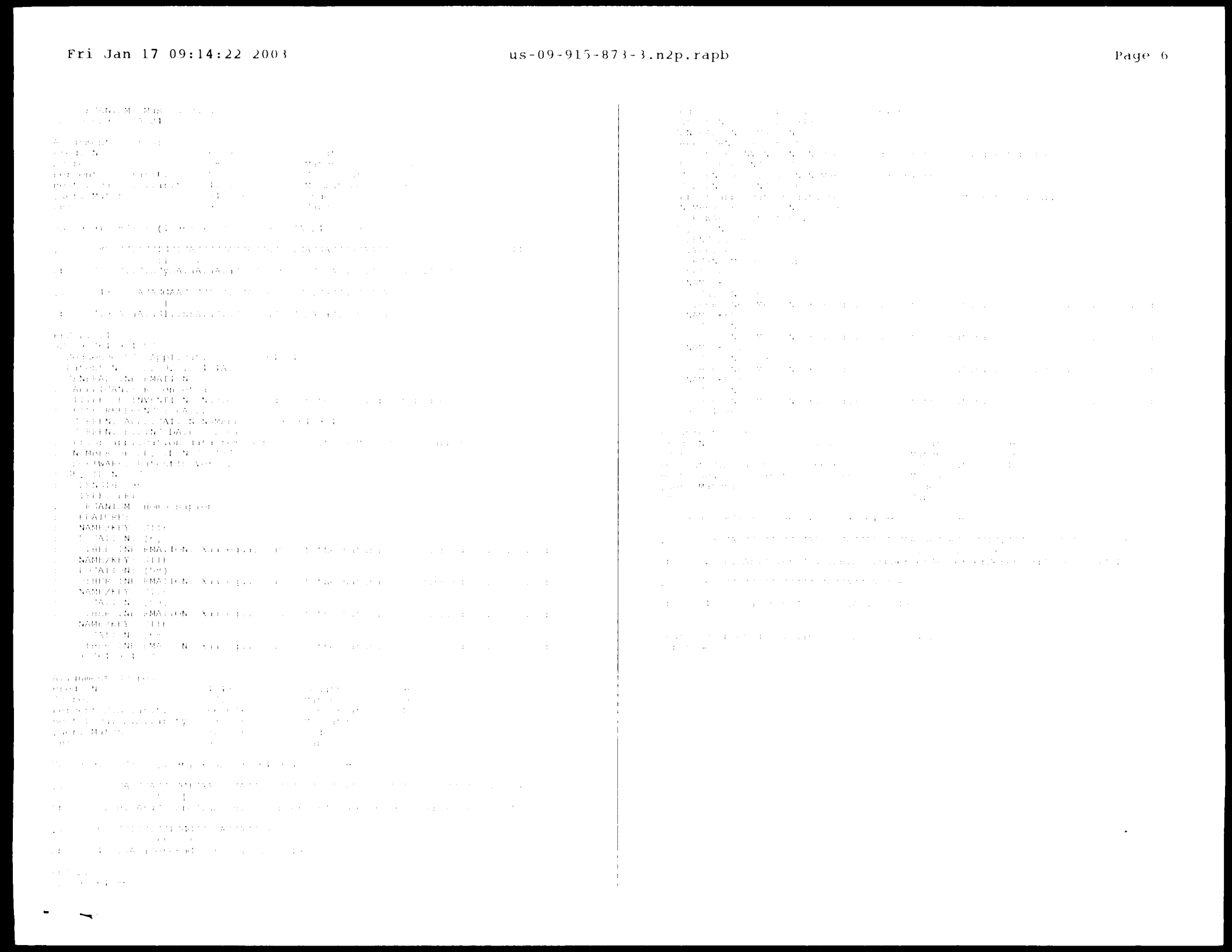
RESULT 1
US-10-057-558-12
; Sequence 12, Application US/10057558
; Publication No. US20020164686A1
; GENERAL INFORMATION:
; AUTHOR(S): Chapman, Kevin
; TITLE OF INVENTION: [illegible]
; APPLICANT: Donovan, Jonathan
; APPLICANT: Pomponi, Monica M.
; APPLICANT: [illegible]
; APPLICANT: Stolarczyk, Shelley
; TITLE OF INVENTION: METHODS FOR MODULATING DNA [illegible]
; TITLE OF INVENTION: [illegible]
; FILE REFERENCE: [illegible]
; CURRENT APPLICATION NUMBER: US/10/057,558
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: [illegible]
; PRIOR FILING DATE: [illegible]
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 32
; TYPE: DNA
; ORGANISM: [illegible]
US-10-057-558-12



```
Best Local Similarity:  96.77%          Mismatches:     1
Query Match:            73.88%          Indels:         1
DB:                     9               Gaps:           1

US-09-915-873-3 (1-96) x US-10-057-558-8 (1-60)

Qy       4 ATGCAGGTCCTGAACACGATGGTCAACAAGCACTTCCTCTCCTTCAACCTC 53
           ||||||||    |||||||||||||||||||||||||||||||||||||||
Db       1 MetGlnVal   AsnThrMetValAsnLysHisPheLeuSerPheAsnVal 37

Qy      54 CTCCTCGGGCTGAGCAGCAACCTCACCGCCGGC 86
           ||||||||||||||||||||||||||||||||
Db      38 LeuLeuGlyLeuSerSerAsnLeuThrAlaGly 60

RESULT 5
US-10-057-558-6
; Sequence 6, Application US/10057558
; Patent No. US20020164585A1
; GENERAL INFORMATION:
;  APPLICANT: Chapman, Sean
;  APPLICANT:  Dawson, William O.
;  APPLICANT:  Donson, Jonathan
;  APPLICANT:  Kumagai, Monto H.
;  APPLICANT:  Lewandowski, Dennis J.
;  APPLICANT:  Lindo, John A.
;  APPLICANT:  Poque, Gregory P.
;  APPLICANT:  Shivprasad, Shailaja
;  TITLE OF INVENTION: METHOD FOR ENHANCING RNA OR PROTEIN
;  TITLE OF INVENTION:  PRODUCTION USING NON NATIVE 5' UNTRANSLATED SEQUENCES IN
;  TITLE OF INVENTION:  RECOMBINANT VIRAL NUCLEIC ACIDS
;  FILE REFERENCE: 008010170508
;  CURRENT APPLICATION NUMBER: US/10/057,558
;  CURRENT FILING DATE:  2002-01-25
;  PRIOR APPLICATION NUMBER: 09/359,299
;  PRIOR FILING DATE: 1999-07-21
;  NUMBER OF SEQ ID NOS: 49
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
;   LENGTH: 31
;   TYPE: PRT
;   ORGANISM: Tobacco mosaic virus
US-10-057-558-6

Alignment Scores:
Pred. No.:              9.76e-68        Length:         31
Score:                  115.00          Matches:        29
Percent Similarity:     93.55%          Conservative:   0
Best Local Similarity:  93.55%          Mismatches:     0
Query Match:            64.61%          Indels:         2
DB:                     9               Gaps:           2

US-09-915-873-3 (1-96) x US-10-057-558-6 (1-31)

Qy       4 ATGCAGGTCCTGAACACGATGGTCAACAAGCACTTCCTCTCCTTCAACCTC 53
           ||||||||||||||||||||    |||||||||    ||||||||||||||||||||||
Db       1 MetGlnValLeuAsnThrMet---AsnLysHis--LeuSerLeuSerValLeuAsnVal 38

Qy      54 CTCCTCGGGCTGAGCAGCAACCTCACCGCCGGC 86
           ||||||||||||||||||||||||||||||||
Db      19 LeuLeuGlyLeuSerSerAsnLeuThrAlaGly 29

RESULT 6
US-10-057-558-2
; Sequence 2, Application US/10057558
; Patent No. US20020164585A1
; GENERAL INFORMATION:
;  APPLICANT: Chapman, Sean
;  APPLICANT:  Dawson, William O.
;  APPLICANT:  Donson, Jonathan
;  APPLICANT:  Kumagai, Monto H.
;  APPLICANT:  Lewandowski, Dennis J.
;  APPLICANT:  Lindo, John A.
;  APPLICANT:  Poque, Gregory P.
;  APPLICANT:  Shivprasad, Shailaja
;  TITLE OF INVENTION: METHOD FOR ENHANCING RNA OR PROTEIN
;  TITLE OF INVENTION:  PRODUCTION USING NON NATIVE 5' UNTRANSLATED SEQUENCES IN
;  TITLE OF INVENTION:  RECOMBINANT VIRAL NUCLEIC ACIDS
;  FILE REFERENCE: 008010170508
;  CURRENT APPLICATION NUMBER: US/10/057,558
;  CURRENT FILING DATE:  2002-01-25
;  PRIOR APPLICATION NUMBER: 09/359,299
;  PRIOR FILING DATE: 1999-07-21
;  NUMBER OF SEQ ID NOS: 49
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 28
;   TYPE: PRT
;   ORGANISM: Tobacco mosaic virus
US-10-057-558-2

Alignment Scores:
Pred. No.:              [illegible]     Length:         28
Score:                  [illegible]     Matches:        25
Percent Similarity:     [illegible]     Conservative:   [illegible]
Best Local Similarity:  [illegible]     Mismatches:     [illegible]
Query Match:            [illegible]     Indels:         [illegible]
DB:                     [illegible]     Gaps:           [illegible]

US-09-915-873-3 (1-96) x US-10-057-558-2 (1-28)

Qy       4 ATGCAGGTCCTGAACACGATGGTCAACAAGCACTTCCTCTCCTTCAACCTC 53
           ||||||||||||||||||||    |||||||||    ||||||||||||||||||||||
Db       1 MetGlnValLeuAsnThrMet---AsnLysHis--LeuSerLeuSerValLeuAsnVal 38

Qy      54 CTCCTCGGGCTGAGCAGCAACCTCACCGCCGGC 86
           ||||||||||||||||||||||||||||||||
Db      19 LeuLeuGlyLeuSerSerAsnLeuThrAlaGly 29

RESULT 7
US-09-[illegible]-[illegible]
; Sequence 17, Application US/[illegible]
; Patent No. US20020164585A1
; GENERAL INFORMATION:
;  APPLICANT: [illegible]
;  TITLE OF INVENTION: [illegible]
;  FILE REFERENCE: [illegible]
;  CURRENT APPLICATION NUMBER: US/[illegible]
;  CURRENT FILING DATE: [illegible]
;  PRIOR APPLICATION NUMBER: [illegible]
;  PRIOR FILING DATE: [illegible]
;  PRIOR APPLICATION NUMBER: [illegible]
;  PRIOR FILING DATE: [illegible]
;  PRIOR APPLICATION NUMBER: [illegible]
;  PRIOR FILING DATE: [illegible]
;  NUMBER OF SEQ ID NOS: [illegible]
;  SOFTWARE: [illegible]
; SEQ ID NO 17
;   LENGTH: 516
;   TYPE: PRT
;   ORGANISM: Unknown [illegible]
;   FEATURE:
;   OTHER INFORMATION: N, [illegible]
;   OTHER INFORMATION: N, [illegible]
US-09-[illegible]-[illegible]

Alignment Scores:
Pred. No.:              [illegible]     Length:         [illegible]
Score:                  [illegible]     Matches:        [illegible]
Percent Similarity:     [illegible]     Conservative:   [illegible]
Best Local Similarity:  [illegible]     Mismatches:     [illegible]
Query Match:            [illegible]     Indels:         [illegible]
DB:                     [illegible]     Gaps:           [illegible]

US-09-915-873-3 (1-96) x US-[illegible] (1-[illegible])

Qy      22 ATGGTCAACAAGCACTTCCTCTCCTTCAACCTCCTCCTCGGGCTGAGCAGCAACCTC 81
```



```
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,812
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,866
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/808,468
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 36100
;   LENGTH: 246
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL035414.23
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.4
;   OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.2
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.6
;   OTHER INFORMATION: EXPRESSED IN MCNP HAPPY, SIGNAL = 1.2
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.4
;   OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1
;   OTHER INFORMATION: SWISSPROT HIT: P26480, EVALUE 2.6e-61
;   OTHER INFORMATION: EST_HUMAN HIT: AL047599.1, EVALUE 4.00e-47
US-09-864-761-36100

Alignment Scores:
Pred. No.:                4.34            Length:          246
Score:                    58.00           Matches:         32
Percent Similarity:       78.26%          Conservative:    6
Best Local Similarity:    52.17%          Mismatches:      5
Query Match:              32.58%          Indels:          0
DB:                       10              Gaps:            0

US-09-915-873-3 (1-96) x US-09-864-761-36100 (1-246)

Qy      26 TCAACAAGCACTTCCTCTCCCTGCTGCTCCACTCTTCCTTCTGGTGATATAAAAG 85
           |||   |||||||:::::     |||||||::|||||||||||.|   :::::::::
Db     102 SerAlaSerThrAlaAlaSerSerProSerAlaSerSerSerSerSerLeuIleSerSer 141

Qy      86 TCAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAA 145
           |||| ||||
Db     142 SerArgSerTrpTrpArgLys 154

RESULT 12
US-09-864-761-11
; Sequence 11, Application US/09864761
; Patent No. US20020127681
; GENERAL INFORMATION:
;  APPLICANT: Goodsell, Allen
;  TITLE OF INVENTION: Polypeptides
;  FILE REFERENCE: PF186.P1.US
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE: 2001-05-23
;  PRIOR APPLICATION NUMBER: US/60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US/60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US/09/808,468
;  PRIOR FILING DATE: 2000-06-30
;  NUMBER OF SEQ ID NOS: 49
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
;   LENGTH: 477
;   TYPE: PRT
;   ORGANISM: Thermococcus litoralis
US-09-864-761-11

Alignment Scores:
Pred. No.:                4.34            Length:          477
Score:                    58.00           Matches:         32
Percent Similarity:       78.26%          Conservative:    6
Best Local Similarity:    52.17%          Mismatches:      5
Query Match:              32.58%          Indels:          0
DB:                       10              Gaps:            0

US-09-915-873-3 (1-96) x US-09-864-761-11 (1-477)

Qy      26 TCAACAAGCACTTCCTCTCCCTGCTGCTCCACTCTTCCTTCTGGTGATATAAAAG 85
           |||   |||||||:::::     |||||||::|||||||||||.|   :::::::::
Db     102 SerAlaSerThrAlaAlaSerSerProSerAlaSerSerSerSerSerLeuIleSerSer 141

Qy      86 TCAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAA 145
           |||| ||||
Db     142 SerArgSerTrpTrpArgLys 154

RESULT 13
US-09-864-761-24
; Sequence 24, Application US/09864761
; Patent No. US20020127681
; GENERAL INFORMATION:
;  APPLICANT: OLSON, ERIC N.
;  APPLICANT: WANG, DA-ZHI
;  TITLE OF INVENTION: MUSCLE-SPECIFIC CALCINEURIN-BINDING PROTEIN
;  TITLE OF INVENTION: NUCLEAR LOCALIZATION SIGNALS
;  FILE REFERENCE: UTSD:P605
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE: 2001-05-23
;  PRIOR APPLICATION NUMBER: 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  NUMBER OF SEQ ID NOS: 56
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 24
;   LENGTH: 1967
;   TYPE: DNA
```



GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: January 17, 2003, 06:54:07 ; Search time 14.0227 Seconds
(without alignments)
346.210 Million cell updates/sec

Title: US-09-915-873-5
Perfect score: 851
Sequence: 1 CDLPQTHSLGSRRTLMLLAQ.....EIMRSFSLSINLQESLFSKE 165

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29122922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/ina/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

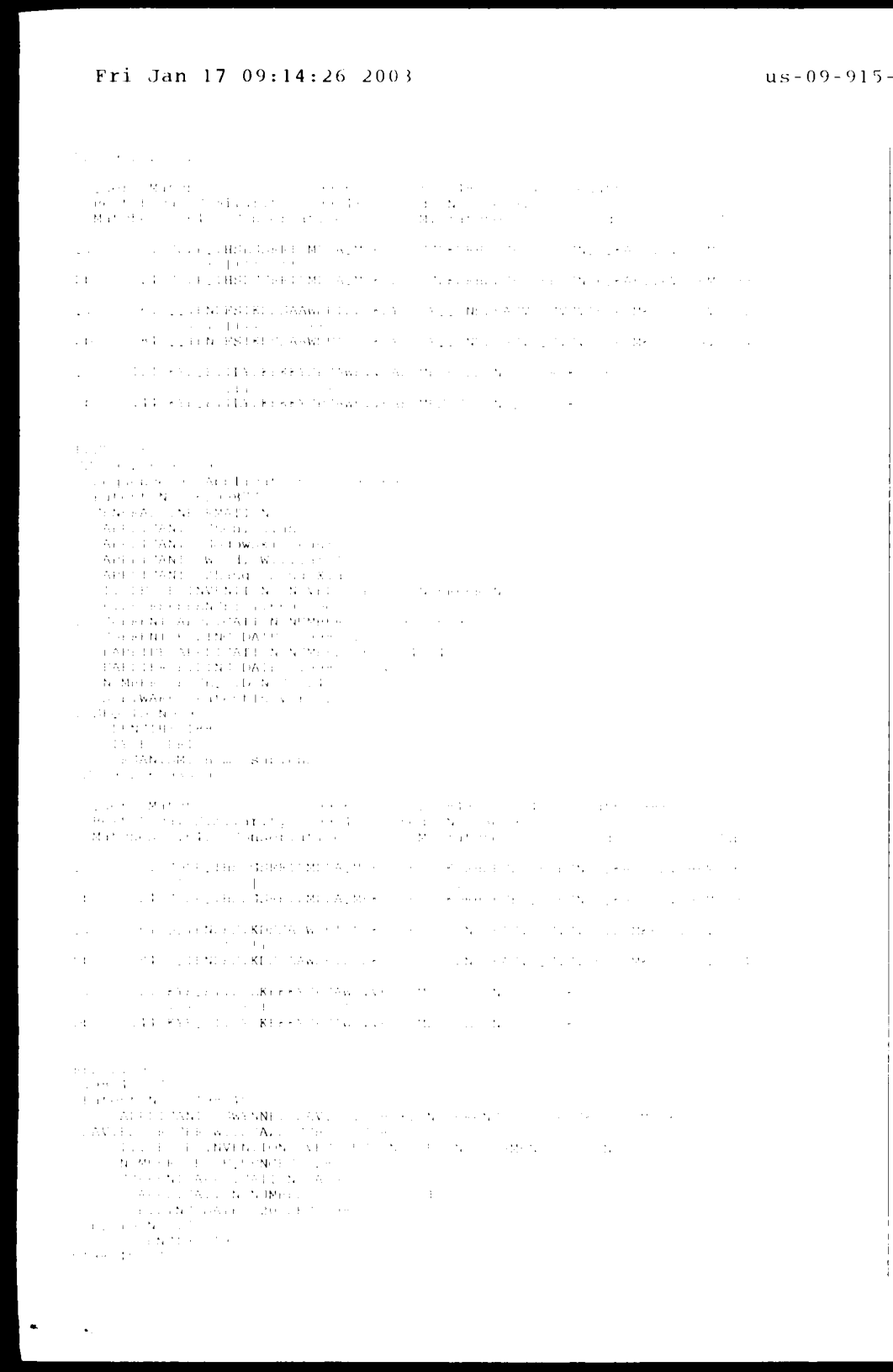
Result No.	Score	% Query Match	Length	DB	ID	Description
1	851	100.0	165	1	US-08-024-188-1	Sequence 1, Appli...
2	851	100.0	165	1	US-07-952-841-1	Sequence 1, Appli...
3	851	100.0	165	1	US-08-356-021-1	Sequence 1, Appli...
4	851	100.0	165	5	PCT-US94-01729-1	Sequence 1, Appli...
5	851	100.0	198	4	US-09-333-348B-11	Sequence 11, Appl...
6	848	99.6	188	4	US-09-206-903A-7	Sequence 7, Appli...
7	848	99.6	198	4	US-09-202-122-7	Sequence 7, Appl...
8	848	99.6	188	4	US-09-206-935-9	Sequence 9, Appli...
9	846	99.4	165	6	5198415-17	Patent No. 5198...
10	845	99.3	165	4	US-09-216-500-1	Sequence 1, Appl...
11	845	99.3	165	4	US-09-730-464-1	Sequence 1, Appl...
12	843	99.1	165	1	US-08-249-671A-5	Sequence 5, Appl...
13	843	99.1	188	1	US-08-249-671A-11	Sequence 11, Appl...
14	840.5	98.8	166	4	US-08-819-238A-1	Sequence 1, Appl...
15	840.5	98.8	166	4	US-09-379-434-1	Sequence 1, Appl...
16	835	98.1	165	6	5210029-3	Patent No. 5210...
17	835	98.1	219	6	5218729-1	Patent No. 5218...
18	834	98.0	188	4	US-09-206-936-9	Sequence 9, Appl...
19	833.5	97.9	166	4	US-08-819-238A-13	Sequence 13, Appl...
20	833.5	97.9	166	4	US-09-379-434-13	Sequence 13, Appl...
21	829.5	97.5	166	1	US-08-362-453-9	Sequence 9, Appli...
22	828.5	97.4	189	1	US-08-026-758-4	Sequence 4, Appl...
23	825.5	97.0	189	1	US-08-026-758-5	Sequence 5, Appli...
24	823.5	96.8	189	6	5519472-7	Patent No. 5519...
25	815	95.8	165	1	US-08-362-453-13	Sequence 13, Appl...
26	810	95.2	165	1	US-09-362-453-14	Sequence 14, Appl...
27	808	94.9	165	6	5256769-2	Patent No. 5256...
28	762	89.5	165	1	US-08-024-188-2	Sequence 2, Appl...
29	762	89.5	165	1	US-08-024-188-4	Sequence 4, Appl...
30	762	89.5	165	1	US-08-024-188-12	Sequence 12, Appl...
31	754	88.6	165	4	PCT-US94-12[illegible]	Patent No. [illegible]
32	748.5	87.9	165	1	US-08-024-188-8	Sequence 8, Appl...
33	746.5	87.7	189	4	US-09-[illegible]-11	Sequence 11, Appl...
34	721.5	84.8	166	4	US-[illegible]-[illegible]-8	Sequence 8, Appl...
35	721.5	84.8	168	4	US-[illegible]-[illegible]-1A	Sequence 1, Appl...
36	721.5	84.8	166	4	US-[illegible]-[illegible]-2B	Sequence 2, Appl...
37	721.5	84.8	166	1	US-08-[illegible]-[illegible]-26	Sequence 26, Appl...
38	721.5	84.8	166	4	US-[illegible]-[illegible]-7	Sequence 7, Appl...
39	721.5	84.8	189	1	US-[illegible]-[illegible]-7	Sequence 7, Appl...
40	721.5	84.8	189	4	US-[illegible]-[illegible]-11	Sequence 11, Appl...
41	721.5	84.8	189	4	US-[illegible]-[illegible]-12	Sequence 12, Appl...
42	719.5	84.5	189	4	US-[illegible]-[illegible]-13	Sequence 13, Appl...
43	716.5	84.2	189	1	US-[illegible]-[illegible]-7	Sequence 7, Appl...
44	714.5	84.0	166	4	US-[illegible]-[illegible]-1B	Sequence 1, Appl...
45	707.5	83.1	189	1	US-[illegible]-[illegible]-1A	Sequence 1, Appl...

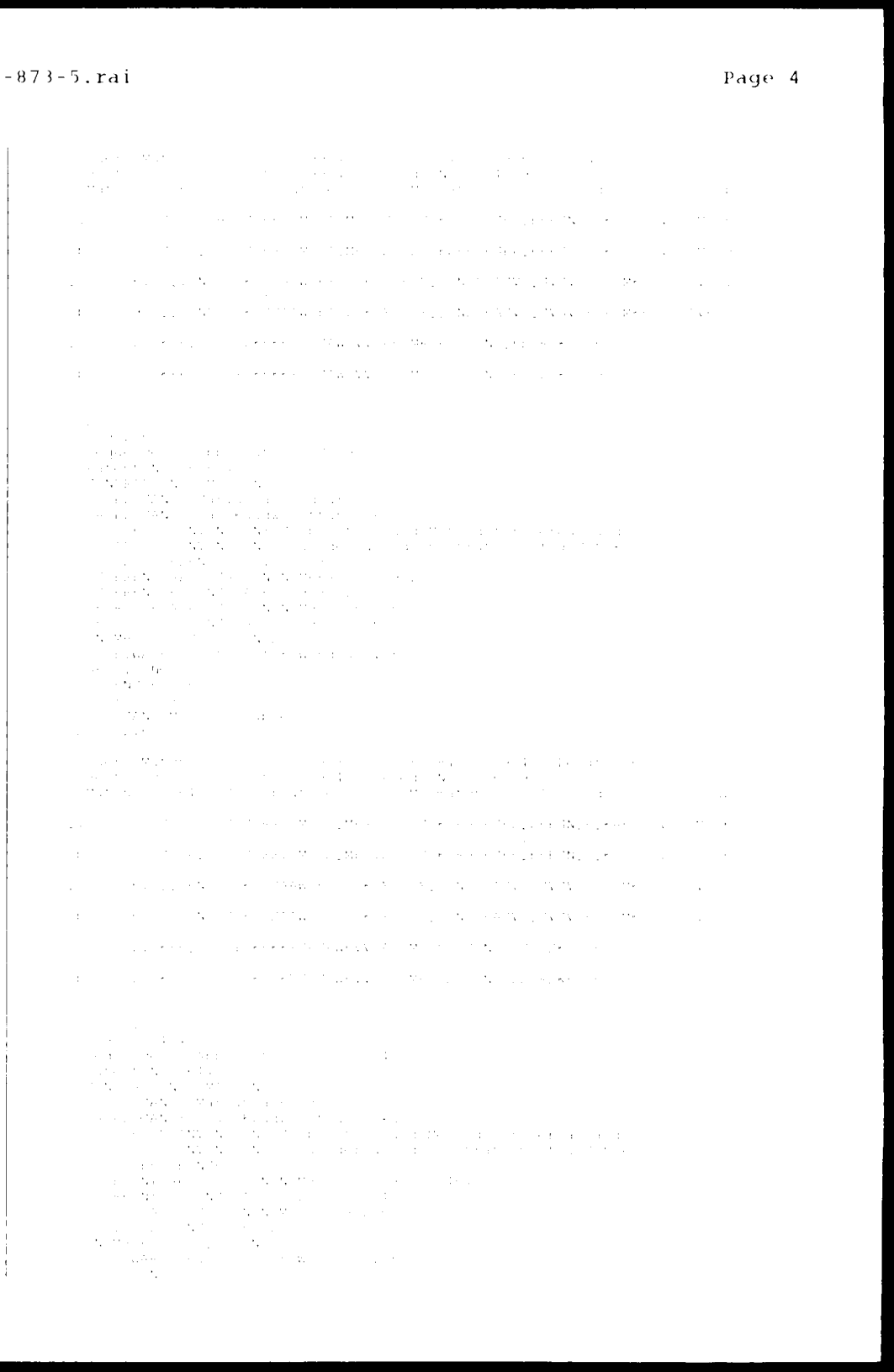
ALIGNMENTS

RESULT 1
US-08-024-188-1
; Sequence 1, Application US/08024188
; Patent No. 5441734
; GENERAL INFORMATION:
; APPLICANT: Reichert, Paul
; APPLICANT: McNemar, Charles
; APPLICANT: Nanabhushan, Narendra
; APPLICANT: Nanabhushan, Tattanahalli...
; APPLICANT: Lindall, Stephen
; APPLICANT: Rose, Alan
; TITLE OF INVENTION: Novel Interferon-Alfa Crystals
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schering-Plough [illegible]
; STREET: One Giralda Farms
; CITY: Madison
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07940
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh [illegible]
; SOFTWARE: Microsoft Word 5.0a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/024,188
; FILING DATE: [illegible]
; CLASSIFICATION: [illegible]
; PRIOR APPLICATION DATA: [illegible]
; ATTORNEY/AGENT INFORMATION:
; NAME: Dodd, Paul L.
; REGISTRATION NUMBER: [illegible]
; REFERENCE/DOCKET NUMBER: [illegible]
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: [illegible]
; TELEFAX: [illegible]
; TELEX: [illegible]
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: [illegible]
; TYPE: AMINO ACID
; TOPOLOGY: Linear
; MOLECULE TYPE: protein
US-08-024-188-1

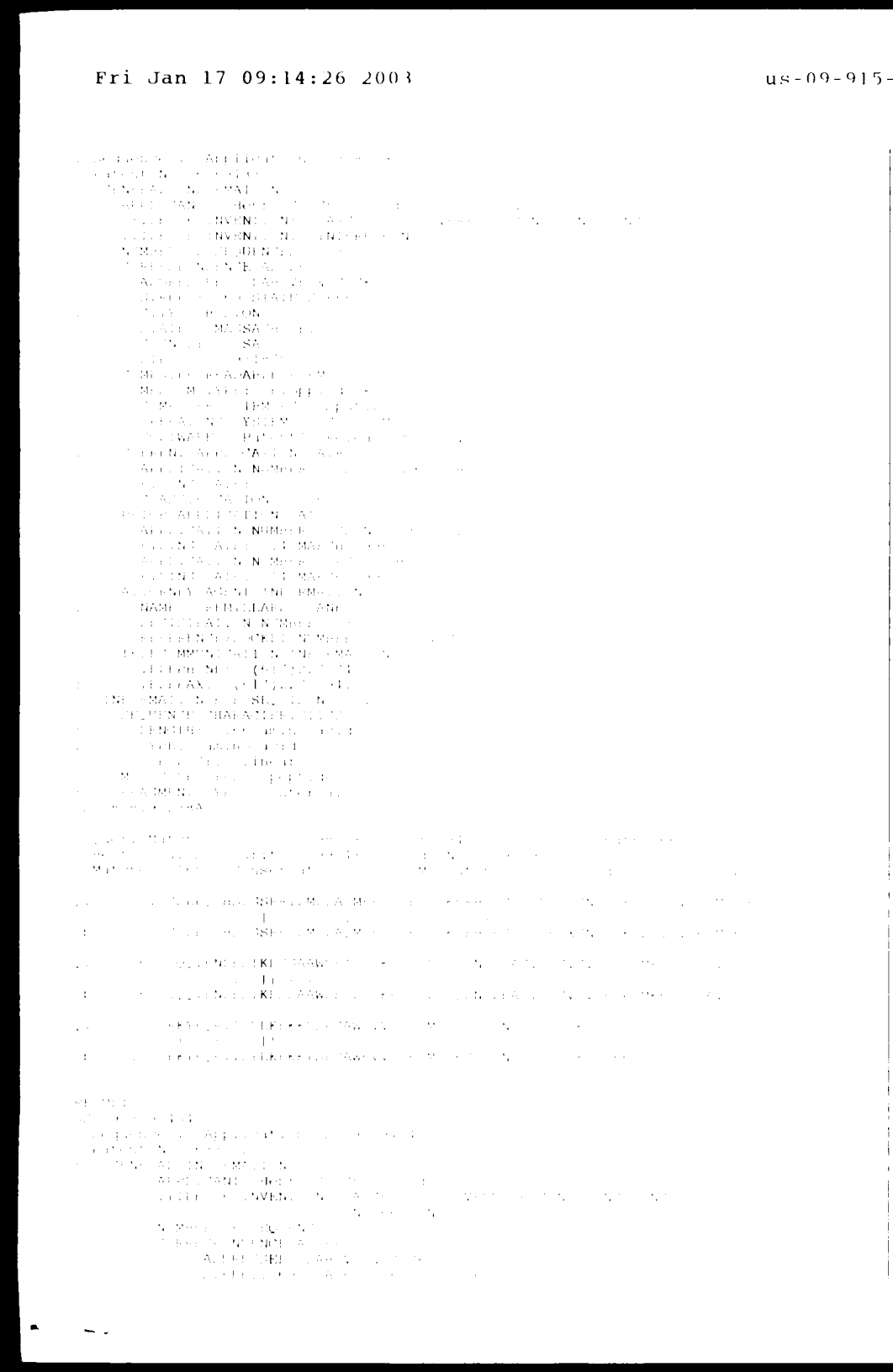
Query Match [illegible]; Score [illegible]; DB [illegible]; Length [illegible];
Best Local Similarity [illegible]; Pred. No. [illegible];
Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

```
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Schering-Plough Corporation
;      STREET:  One Giralda Farms
;      CITY:  Madison
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07940
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  Apple Macintosh
;      OPERATING SYSTEM:  Macintosh 6.0.8
;      SOFTWARE:  Microsoft Word 5.15
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/01729
;      FILING DATE:  Herewith
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: US/08/024,330
;      FILING DATE:  25-FEB-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lunn, Paul G.
;      REGISTRATION NUMBER:  32,743
;      REFERENCE/DOCKET NUMBER:  DBC287K
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201 822 7259
;      TELEFAX:  201 822 7039
;      TELEX:  219165
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  165 amino acid residues
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
PCT-US94-01729-1

  Query Match             100.0%;  Score 851;  DB 5;  Length 165;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-91;
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 CDLPQTHSLGSRRTLMLLAQMRRISLFSCLKDRHDFGFPQEEFGNQFQKAETIPVLHEMI 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 CDLPQTHSLGSRRTLMLLAQMRRISLFSCLKDRHDFGFPQEEFGNQFQKAETIPVLHEMI 60

Qy       61 QQIFNLFSTKDSSAAWDETLLDKFYTELYQQLNDLEACVIQGVGVTETPLMKEDSILAVR 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 QQIFNLFSTKDSSAAWDETLLDKFYTELYQQLNDLEACVIQGVGVTETPLMKEDSILAVR 120

Qy      121 KYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTNLQESLRSKE 165
            |||||||||||||||||||||||||||||||||||||||||||||
Db      121 KYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTNLQESLRSKE 165


RESULT 5
US-09-333-348B-11
; Sequence 11, Application US/09333348B
; Patent No. 6319691
; GENERAL INFORMATION:
;  APPLICANT: Pang, Danny Z.
;  TITLE OF INVENTION: Fusion Proteins Comprising IFN-alpha2b and IL-alpha1
;  FILE REFERENCE:
;  CURRENT APPLICATION NUMBER: US/09/333,348B
;  CURRENT FILING DATE:  1999-06-15
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: Patentin ver. 2.0
; SEQ ID NO 11
;   LENGTH: 198
;   TYPE: PRT
;   ORGANISM: Human
US-09-333-348B-11

  Query Match             100.0%;  Score 851;  DB 4;  Length 198;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-91;
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 CDLPQTHSLGSRRTLMLLAQMRRISLFSCLKDRHDFGFPQEEFGNQFQKAETIPVLHEMI 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       34 CDLPQTHSLGSRRTLMLLAQMRRISLFSCLKDRHDFGFPQEEFGNQFQKAETIPVLHEMI 93

Qy       61 QQIFNLFSTKDSSAAWDETLLDKFYTELYQQLNDLEACVIQGVGVTETPLMKEDSILAVR 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       94 QQIFNLFSTKDSSAAWDETLLDKFYTELYQQLNDLEACVIQGVGVTETPLMKEDSILAVR 153

Qy      121 KYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTNLQESLRSKE 165
            |||||||||||||||||||||||||||||||||||||||||||||
Db      154 KYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTNLQESLRSKE 198


RESULT 7
US-09-202-122-7
; Sequence 7, Application US/09202122
; Patent No. 6200557
; GENERAL INFORMATION:
;  APPLICANT:  Chen, Jian
;  APPLICANT:  Dobrzanski, Pet
;  APPLICANT:  Wood, William I.
;  APPLICANT:  Zhang, Dong-Xin
;  TITLE OF INVENTION:  HUMAN INTERFERON ALPHA RECEPTOR
;  FILE REFERENCE:  P1224R2 (EP-01)
;  CURRENT APPLICATION NUMBER:  US/09/202,122
;  CURRENT FILING DATE:  1998-12-07
;  PRIOR APPLICATION NUMBER:  PCT/US97/10144
;  PRIOR FILING DATE:  1997-06-13
;  NUMBER OF SEQ ID NOS:  12
; SEQ ID NO 7
;   LENGTH: 198
;   TYPE: PRT
;   ORGANISM: Homo sapiens
```





```
;   LENGTH: 165
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-730-464-1

  Query Match             99.4%;  Score 845;  DB 4;  Length 165;
  Best Local Similarity   99.4%;  Pred. No. 5.6e-91;
  Matches  164;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CDLPQTHSLGSRRTLMLLAQMRRISLFSCLKDRHDFGFPQEEFGNQFQKAETIPVLHEM  60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CDLPQTHSLGSRRTLMLLAQMRRISLFSCLKDRHDFGFPQEEFGNQFQKAETIPVLHEM  60

Qy      61 QQIFNLFSTKDSSAAWDETLLDKFYTELYQQLNDLEACVIQGVGVTETPLMKEDSILAVR  120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 QQIFNLFSTKDSSAAWDETLLDKFYTELYQQLNDLEACVIQGVGVTETPLMKEDSILAVR  120

Qy     121 KYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTNLQERLRRK  165
           |||||||||||| ||||||||||||||||||||||||||||||||
Db     121 KYFQRIILYLKEKNYSPCAWEVVRAEIMRSFSLSTNLQERLRRK  165


RESULT 12
US-08-249-671A-5
; Sequence 5, Application US/08249671A
; Patent No. 5710027
;  GENERAL INFORMATION:
;    APPLICANT:  Hauptmann, R.
;    APPLICANT:  Falkner, E.
;    APPLICANT:  Bodo, G.
;    APPLICANT:  Voa, T.
;    APPLICANT:  Maurer-Fogy, I.
;    TITLE OF INVENTION:  Process for Preparing and Purifying
;    TITLE OF INVENTION:  alpha-Interferon
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C.
;      STREET:  1100 New York Avenue, Suite 600
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/249,671A
;      FILING DATE:  26-MAY-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Esmond, Robert W.
;      REGISTRATION NUMBER:  32,893
;      REFERENCE/DOCKET NUMBER:  0652.1350000
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  165 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-249-671A-5

  Query Match             99.1%;  Score 843;  DB 1;  Length 165;
  Best Local Similarity   99.4%;  Pred. No. 9.6e-91;
  Matches  164;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CDLPQTHSLGSRRTLMLLAQMRRISLFSCLKDRHDFGFPQEEFGNQFQKAETIPVLHEM  60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CDLPQTHSLGSRRTLMLLAQMRRISLFSCLKDRHDFGFPQEEFGNQFQKAETIPVLHEM  60

Qy      61 QQIFNLFSTKDSSAAWDETLLDKFYTELYQQLNDLEACVIQGVGVTETPLMKEDSILAVR  120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 QQIFNLFSTKDSSAAWDETLLDKFYTELYQQLNDLEACVIQGVGVTETPLMKEDSILAVR  120

Qy     121 KYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTNLQERLRRK  165
           |||||||||||| ||||||||||||||||||||||||||||||||
Db     121 KYFQRIILYLKEKNYSPCAWEVVRAEIMRSFSLSTNLQERLRRK  165


RESULT 13
US-08-249-671A-11
; Sequence 11, Application US/08249671A
; Patent No. 5710027
;  GENERAL INFORMATION:
;    APPLICANT:  Hauptmann, R.
;    APPLICANT:  Falkner, E.
;    APPLICANT:  Bodo, G.
;    APPLICANT:  Voa, T.
;    APPLICANT:  Maurer-Fogy, I.
;    TITLE OF INVENTION:  Process for Preparing and Purifying
;    TITLE OF INVENTION:  alpha-Interferon
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C.
;      STREET:  1100 New York Avenue, Suite 600
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/249,671A
;      FILING DATE:  26-MAY-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Esmond, Robert W.
;      REGISTRATION NUMBER:  32,893
;      REFERENCE/DOCKET NUMBER:  0652.1350000
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  166 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-249-671A-11

  Query Match             99.1%;  Score 843;  DB 1;  Length 166;
  Best Local Similarity   99.4%;  Pred. No. 9.6e-91;
  Matches  164;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CDLPQTHSLGSRRTLMLLAQMRRISLFSCLKDRHDFGFPQEEFGNQFQKAETIPVLHEM  60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      24 CDLPQTHSLGSRRTLMLLAQMRRISLFSCLKDRHDFGFPQEEFGNQFQKAETIPVLHEM  83

Qy      61 QQIFNLFSTKDSSAAWDETLLDKFYTELYQQLNDLEACVIQGVGVTETPLMKEDSILAVR  120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      84 QQIFNLFSTKDSSAAWDETLLDKFYTELYQQLNDLEACVIQGVGVTETPLMKEDSILAVR  143

Qy     121 KYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTNLQERLRRK  165
           |||||||||||| ||||||||||||||||||||||||||||||||
Db     144 KYFQRIILYLKEKNYSPCAWEVVRAEIMRSFSLSTNLQERLRRK  188


RESULT 14
US-08-915-873-5
```



GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: January 17, 2003, 06:54:46 ; Search time 1.2344 Seconds
(without alignments)
918.960 Million cell updates/sec

Title: US-09-915-873-5
Perfect score: 851
Sequence: 1 CDLPQTHSLGSRRTLMLLAQ.....EIMRSFSLSTNLQESLRSKE 165

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 120991 seqs, 19878514 residues

Total number of hits satisfying chosen parameters: 120991

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	851	100.0	188	9	US-10-087-325-2	Sequence 2, Appl
2	848	99.6	165	10	US-09-977-034-9	Sequence 9, Appli
3	848	99.6	188	9	US-09-919-622A-7	Sequence 7, Appl
4	728.5	85.6	166	10	US-09-977-034-11	Sequence 11, App
5	721.5	84.8	166	10	US-09-559-671A-80	Sequence 80, App
6	721.5	84.8	166	10	US-09-977-034-12	Sequence 12, App
7	717.5	84.3	166	10	US-09-559-671A-75	Sequence 75, App
8	707.5	83.1	170	10	US-09-977-034-16	Sequence 16, App
9	703.5	82.7	166	10	US-09-559-671A-76	Sequence 76, App
10	703.5	82.7	166	10	US-09-559-671A-79	Sequence 79, App
11	698.5	82.1	166	10	US-09-977-034-2	Sequence 2, Appl
12	694.5	81.6	166	10	US-09-559-671A-79	Sequence 79, App
13	694.5	81.6	166	10	US-09-977-034-10	Sequence 10, App
14	694.5	81.6	166	10	US-09-977-034-18	Sequence 18, App
15	692.5	81.4	166	10	US-09-977-034-19	Sequence 19, App
16	691.5	81.3	166	10	US-09-559-671A-86	Sequence 86, App
17	691.5	81.3	166	10	US-09-977-034-17	Sequence 17, App
18	690.5	81.1	166	10	US-09-977-034-15	Sequence 15, App
19	690.5	81.1	189	10	US-09-919-497-73	Sequence 73, App
20	688.5	80.9	166	9	US-09-915-873A-85	Sequence 85, App
21	688.5	80.9	166	10	US-09-559-671A-85	Sequence 85, App
22	688.5	80.9	166	10	US-09-977-034-8	Sequence 8, Appl
23	688.5	80.9	166	9	US-09-915-873A-8	Sequence 8, Appl
24	688.5	80.9	189	9	US-09-919-622A-9	Sequence 9, Appl
25	684.5	80.4	166	10	US-09-559-671A-77	Sequence 77, App
26	684.5	80.4	166	10	US-09-559-671A-82	Sequence 82, App
27	684.5	80.4	166	10	US-09-559-671A-84	Sequence 84, App
28	684.5	80.4	166	10	US-09-559-671A-85	Sequence 85, App
29	679.5	79.9	166	10	US-09-559-671A-81	Sequence 81, App
30	679.5	79.9	166	10	US-09-977-034-13	Sequence 13, App
31	664.5	78.1	166	10	US-09-977-034-14	Sequence 14, App
32	660.5	77.6	201	9	US-09-919-497-72	Sequence 72, App
33	552.5	64.9	166	10	US-09-977-034-25	Sequence 25, App
34	547.5	64.3	166	10	US-09-977-034-22	Sequence 22, App
35	526.5	61.9	167	10	US-09-977-034-27	Sequence 27, App
36	524.5	61.6	167	10	US-09-977-034-23	Sequence 23, App
37	524.5	61.6	167	10	US-09-977-034-24	Sequence 24, App
38	521.5	61.3	166	10	US-09-977-034-26	Sequence 26, App
39	508.5	59.7	166	10	US-09-977-034-28	Sequence 28, App
40	505.5	59.4	172	10	US-09-977-034-21	Sequence 21, App
41	505.5	59.4	196	9	US-09-915-873A-6	Sequence 6, Appl
42	497	58.4	172	10	US-09-977-034-20	Sequence 20, App
43	451.5	53.1	172	10	US-09-977-034-31	Sequence 31, App
44	451.5	53.1	196	10	US-09-559-671A-82	Sequence 82, App
45	438.5	51.5	172	10	US-09-977-034-30	Sequence 30, App

ALIGNMENTS

RESULT 1
US-10-087-325-2
; Sequence 2, Application US/10087325
; Patent No. US20020061484
; GENERAL INFORMATION:
; APPLICANT: Esmon, John Todd
; APPLICANT: Esmon, Charles T.
; TITLE OF INVENTION: ACTIVATED PROTEIN C VARIANTS AND POLYPEPTIDES
; FILE REFERENCE: 021375/0010
; CURRENT APPLICATION NUMBER: US/10/087,325
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 60/273,205
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-087-325-2

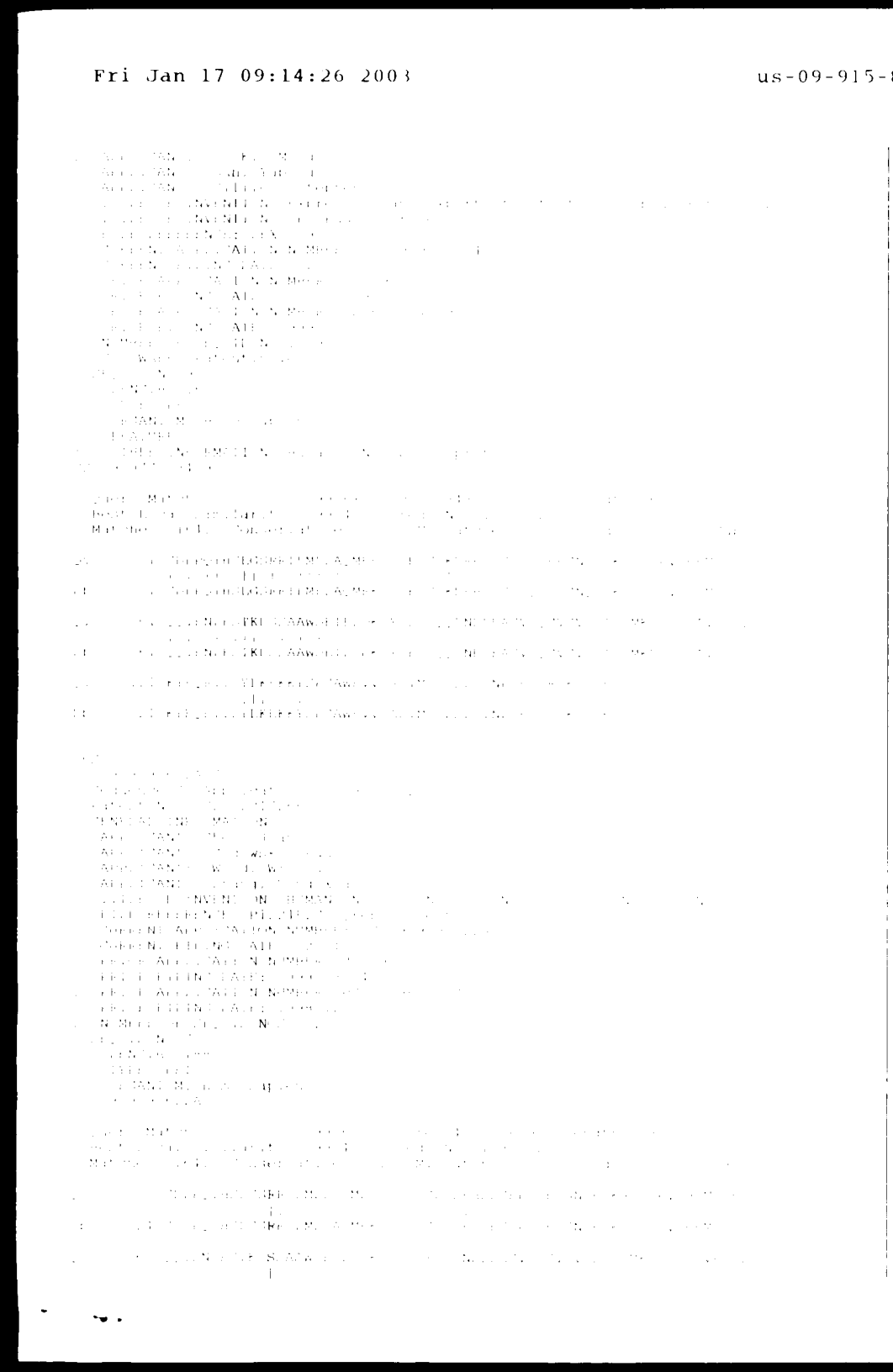
Query Match 100.0%; Score 851; DB 9; Length 188;
Best Local Similarity 100.0%; Pred. No. 4.6e-85;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

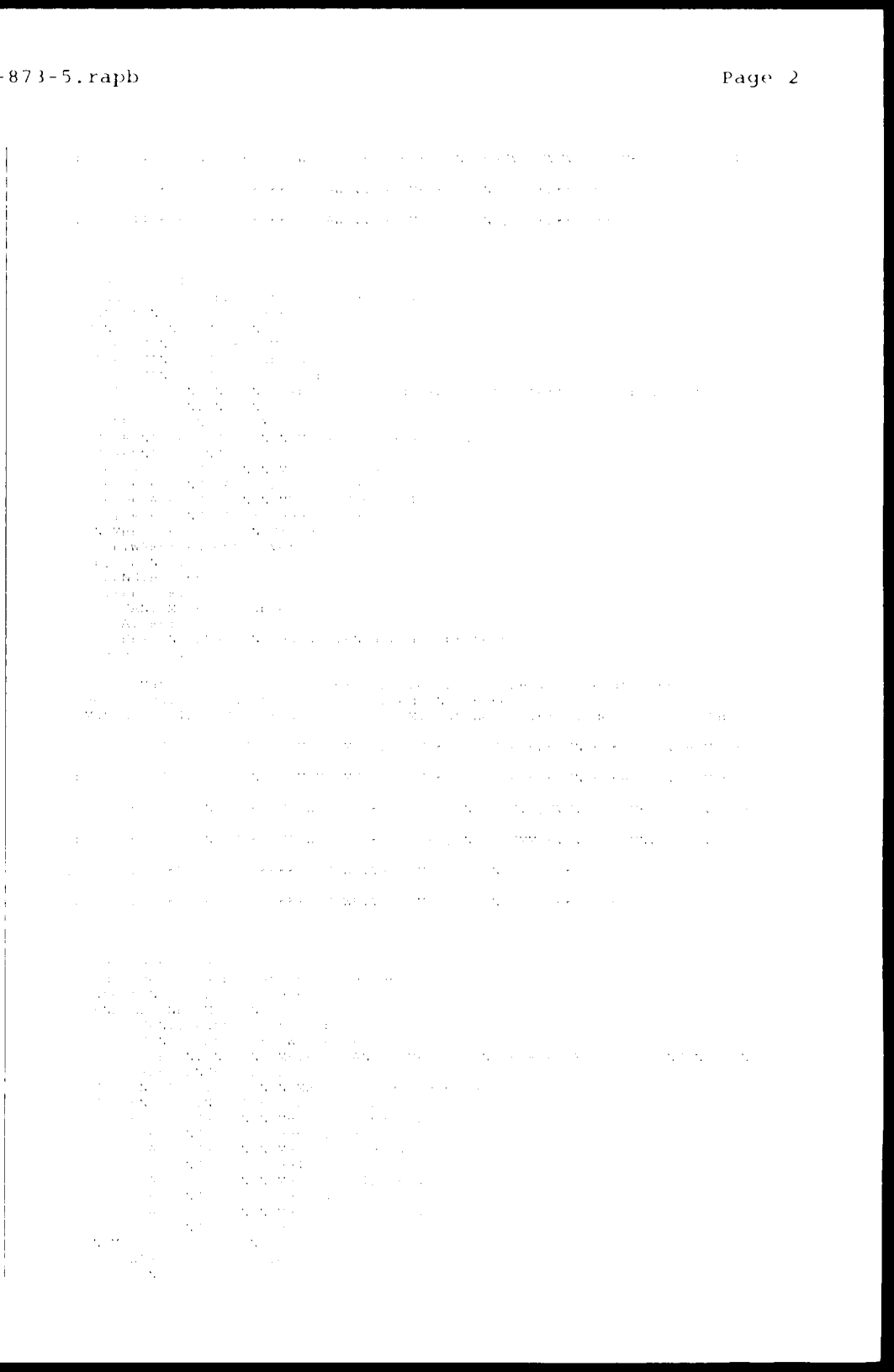
Qy 1 CDLPQTHSLGSRRTLMLLAQMRKISPFSCLKDRHDFGFPQEEFDGNQFQKAPAISVLHE 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 24 CDLPQTHSLGSRRTLMLLAQMRKISPFSCLKDRHDFGFPQEEFDGNQFQKAPAISVLHE 83

Qy 61 LTQQIFNLFSTKDSSAAWDETLLDKFYIELYQQLNDLEACVIQGVGVTETPLMKEDSILAV 120
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 84 LTQQIFNLFSTKDSSAAWDETLLDKFYIELYQQLNDLEACVIQGVGVTETPLMKEDSILAV 143

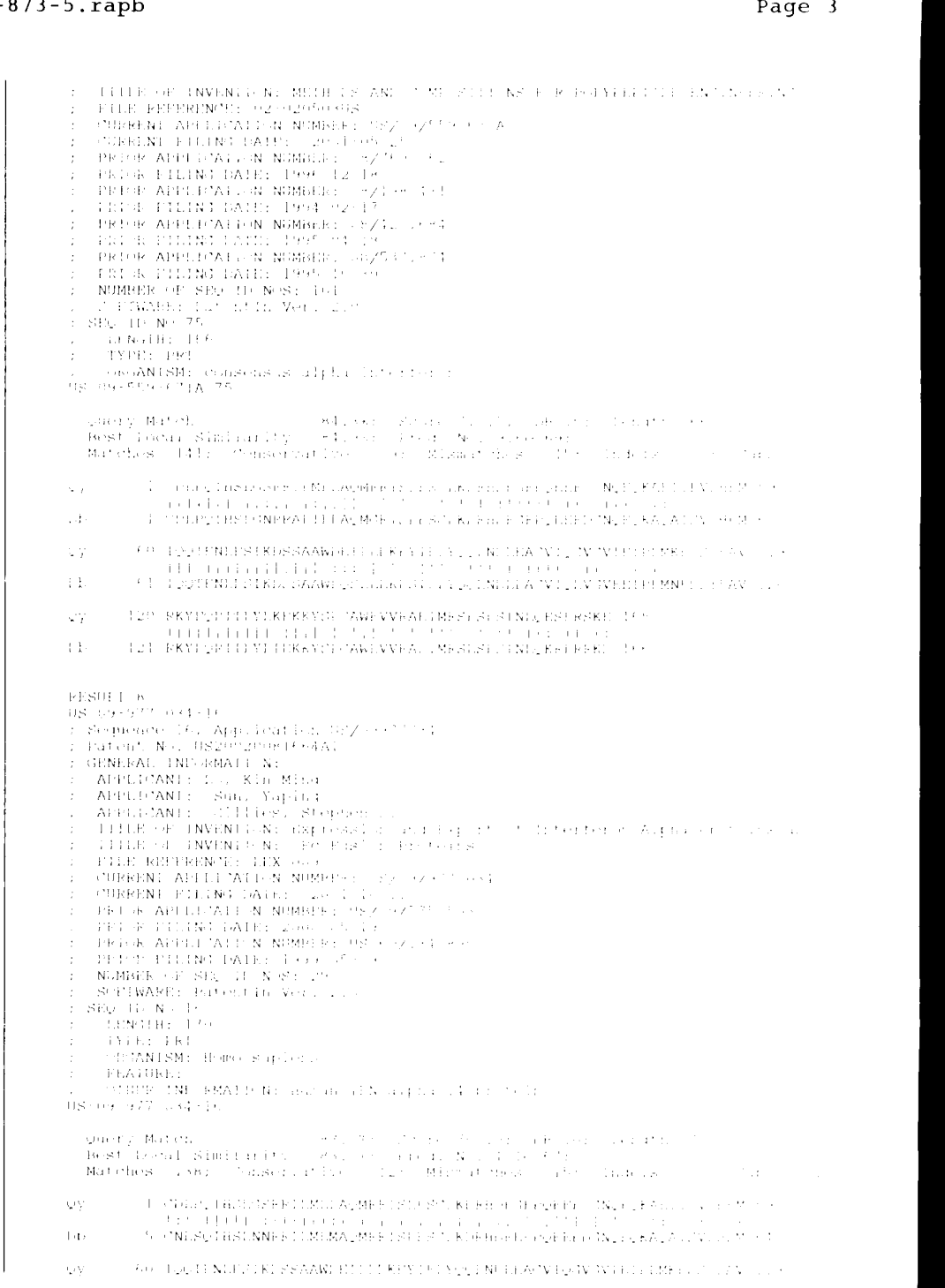
Qy 121 RKYFQRITLYLTEKKYSPCAWEVVRAEIMRSFSLSTNLQESLRSKE 165
|||||||||||||||||||||||||||||||||||||||||||||
Db 144 RKYFQRITLYLTEKKYSPCAWEVVRAEIMRSFSLSTNLQESLRSKE 188

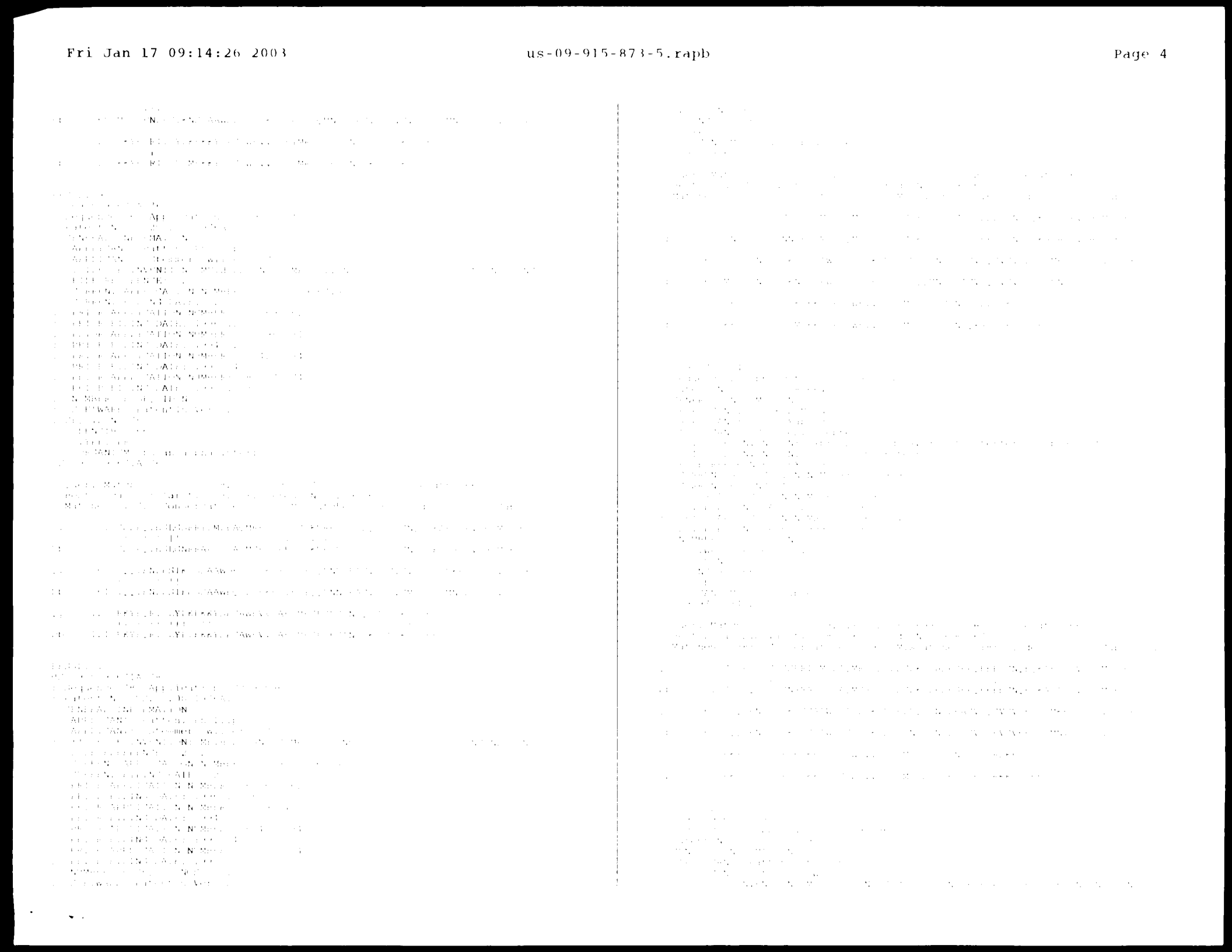
RESULT 2
US-09-977-034-9
; Sequence 9, Application US/09977034
; Patent No. US20020061484A1
; GENERAL INFORMATION:



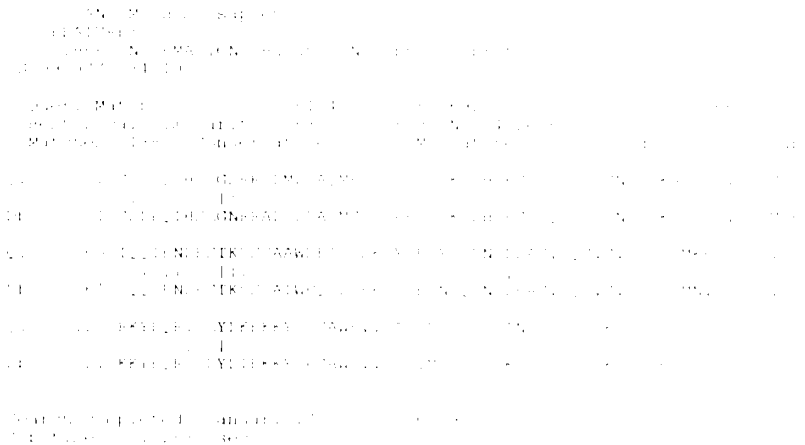


```
;   LENGTH: 166
;   TYPE: PRT
;   ORGANISM: human alpha interferon
US-09-559-671A-80

  Query Match             84.8%;  Score 721.5;  DB 10;  Length 166;
  Best Local Similarity   86.7%;  Pred. No. 3.6e-69;
  Matches  144;  Conservative    3;  Mismatches   18;  Indels    1;  Gaps    1;

Qy       1 CDLPQTHSLGSRRTLMLLAQMRRISLFSCLKDRHDFGFPQEEFGNQFQKAETIPVLHEM 60
           ||||||||| |||||||||||||||||||||||| |||||| ||||||| | ||||
Db       1 CDLPQTHSLGHRRTMMLLAQMRRISLFSCLKDRHDFRFPQEEFDGNQFQKAEAISVLHEV 60

Qy      61 IQQTFNLFSTKDSSAAWDETLLDKFYTELYQQLNDLEACVIQGVGVTETPLMKEDSILAV 120
           ||| |||||||||| |||| |||| |||||||||||||||:| | |  |||| ||||||
Db      61 IQQTFNLFSTKDSSVAWDERLLDKLYTELYQQLNDLEACVMQEVWVGGTPLMNEDSILAV 120

Qy     121 RKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTNLQERLRRKE 166
           |||||||||| ||||||||||||||||||||| | |||| || ||
Db     121 RKYFQRITLYLTEKKYSPCAWEVVRAEIMRSFSLSTNLQERLRRKE 166


RESULT 6
US-09-977-034-12
; Sequence 12, Application US/09977034
; Patent No. US20020081664A1
; GENERAL INFORMATION:
;  APPLICANT: Lo, Kin-Ming
;  APPLICANT:  Sun, Yaping
;  APPLICANT:  Gillies, Stephen D.
;  TITLE OF INVENTION: Expression and Export of Interferon Alpha Proteins as
;  TITLE OF INVENTION:  Fc Fusion Proteins
;  FILE REFERENCE: LEX-009
;  CURRENT APPLICATION NUMBER: US/09/977,034
;  CURRENT FILING DATE:  2001-10-11
;  PRIOR APPLICATION NUMBER: US/09/575,503
;  PRIOR FILING DATE: 2000-05-19
;  PRIOR APPLICATION NUMBER: US 60/134,895
;  PRIOR FILING DATE: 1999-05-19
;  NUMBER OF SEQ ID NOS: 29
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 12
;   LENGTH: 166
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: HUman IFN alpha-6 protein
US-09-977-034-12

  Query Match             84.8%;  Score 721.5;  DB 10;  Length 166;
  Best Local Similarity   86.7%;  Pred. No. 3.6e-69;
  Matches  144;  Conservative    3;  Mismatches   18;  Indels    1;  Gaps    1;

Qy       1 CDLPQTHSLGSRRTLMLLAQMRRISLFSCLKDRHDFGFPQEEFGNQFQKAETIPVLHEM 60
           ||||||||| |||||||||||||||||||||||| |||||| ||||||| | ||||
Db       1 CDLPQTHSLGHRRTMMLLAQMRRISLFSCLKDRHDFRFPQEEFDGNQFQKAEAISVLHEV 60

Qy      61 IQQTFNLFSTKDSSAAWDETLLDKFYTELYQQLNDLEACVIQGVGVTETPLMKEDSILAV 120
           ||| |||||||||| |||| |||| |||||||||||||||:| | |  |||| ||||||
Db      61 IQQTFNLFSTKDSSVAWDERLLDKLYTELYQQLNDLEACVMQEVWVGGTPLMNEDSILAV 120

Qy     121 RKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTNLQERLRRKE 166
           |||||||||| ||||||||||||||||||||| | |||| || ||
Db     121 RKYFQRITLYLTEKKYSPCAWEVVRAEIMRSFSLSTNLQERLRRKE 166


RESULT 7
US-09-559-671A-75
; Sequence 75, Application US/09559671A
; Patent No. US20020051976A1
; GENERAL INFORMATION:
;  APPLICANT: Patten, Phillip
;  APPLICANT:  Stemmer, Willem P.C.
```


; TITLE OF INVENTION: MUTEIN AN
; FILE REFERENCE: 02-1050-US
; CURRENT APPLICATION NUMBER: US/
; CURRENT FILING DATE: 2001-10-
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE: 1999-12-
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE: 1994-12-
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE: 1993-12-
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE: 1993-12-
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 75
;   LENGTH: 166
;   TYPE: PRT
;   ORGANISM: Consensus alpha
US-09-559-671A-75
Query Match
Best Local Similarity
Matches
RESULT 8
US-09-977-034-10
; Sequence 10, Application US/
; Patent No. US20020081664A1
; GENERAL INFORMATION:
; APPLICANT: Lo, Kin-Ming
; APPLICANT: Sun, Yaping
; APPLICANT: Gillies, Stephen
; TITLE OF INVENTION: Expression
; TITLE OF INVENTION: Fc Fusion
; FILE REFERENCE: LEX-
; CURRENT APPLICATION NUMBER:
; CURRENT FILING DATE:
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver.
; SEQ ID NO 10
;   LENGTH:
;   TYPE:
;   ORGANISM:
;   FEATURE:
;   OTHER INFORMATION:
US-09-977-034-10
Query Match
Best Local Similarity
Matches
RESULT 9
US-09-559-671A-77
; Sequence 77, Application US/09559671A
; Patent No. US20020051976A1
; GENERAL INFORMATION:
; APPLICANT: Patten, Phillip
; APPLICANT: Stemmer, Willem P.C.
; TITLE OF INVENTION: MUTEIN AN
; FILE REFERENCE: 02-1050-US
; CURRENT APPLICATION NUMBER: US/
; CURRENT FILING DATE: 2001-10-
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE: 1999-12-
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE: 1994-12-
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE: 1993-12-
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE: 1993-12-
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 77
;   LENGTH: 166
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: Human IFN alpha-A protein
US-09-559-671A-77
Query Match
Best Local Similarity
Matches
```
```



```
;  FILE REFERENCE: 02-0205U3US
;  CURRENT APPLICATION NUMBER: US/09/559,671A
;  CURRENT FILING DATE:  2001-05-25
;  PRIOR APPLICATION NUMBER: 08/769,062
;  PRIOR FILING DATE: 1996-12-18
;  PRIOR APPLICATION NUMBER: 08/198,441
;  PRIOR FILING DATE: 1994-02-17
;  PRIOR APPLICATION NUMBER: 08/425,684
;  PRIOR FILING DATE: 1995-04-18
;  PRIOR APPLICATION NUMBER: 08/537,824
;  PRIOR FILING DATE: 1995-10-30
;  NUMBER OF SEQ ID NOS: 101
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 79
;   LENGTH: 166
;   TYPE: PRT
;   ORGANISM: human alpha interferon
US-09-559-671A-79

  Query Match             81.6%;  Score 694.5;  DB 10;  Length 166;
  Best Local Similarity   81.9%;  Pred. No. 2.6e-66;
  Matches  136;  Conservative   12;  Mismatches   17;  Indels    1;  Gaps    1;

Qy       1 CDLPQTHSLGSRRTLMLLAQMRRISLFSCLKDRHDFGFPQEEFGNQFQKAETIPVLHEM 60
           ||||||||||:|| |:|||| ||| |||||||||||||||||| |||||| | | ||||
Db       1 CDLPQTHSLGNRRALILLAQMGRISHFSCLKDRHDFRFPEEEFDGHQFQKTQAISVLHEM 60

Qy      61 IQQTFNLFSTKDSSAAWDETLLDKFYTELYQQLNDLEACVIQGVGVTETPLMKEDSILAV 120
           ||| |||||:|||||:::|:|:| |||||||||||||||| ||| |||||  ||||||
Db      61 IQQTFNIFSTEDSSAAWIQSLLPKFSTELYEQLNDLEACVIQEVGVEETPLMNVDSILAV 120

Qy     121 RKYFQRITLYLTEKKYSPCAWEVVRAEIMRSLSFSTNLQKRLRRKD 166
           |||||||||| |||||||||||||||||| | ||||| || |:
Db     121 RKYFQRITLYLTEKKYSPCAWEVVRAEIMRSLSFSTNLQKRLRRKD 166


RESULT 13
US-09-977-034-10
; Sequence 10, Application US/09977034
; Patent No. US20020081664A1
; GENERAL INFORMATION:
;  APPLICANT: Lo, Kin-Ming
;  APPLICANT:  Sun, Yaping
;  APPLICANT:  Gillies, Stephen D.
;  TITLE OF INVENTION: Expression and Export of Interferon-Alpha Proteins as
;  TITLE OF INVENTION:  Fc Fusion Proteins
;  FILE REFERENCE: LEX-009
;  CURRENT APPLICATION NUMBER: US/09/977,034
;  CURRENT FILING DATE:  2001-10-11
;  PRIOR APPLICATION NUMBER: US/09/575,583
;  PRIOR FILING DATE: 2000-05-19
;  PRIOR APPLICATION NUMBER: US 60/134,895
;  PRIOR FILING DATE: 1999-05-19
;  NUMBER OF SEQ ID NOS: 29
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
;   LENGTH: 166
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: Human IFN alpha-4 protein
US-09-977-034-10

  Query Match             81.6%;  Score 694.5;  DB 10;  Length 166;
  Best Local Similarity   81.9%;  Pred. No. 2.6e-66;
  Matches  136;  Conservative   12;  Mismatches   17;  Indels    1;  Gaps    1;

Qy       1 CDLPQTHSLGSRRTLMLLAQMRRISLFSCLKDRHDFGFPQEEFGNQFQKAETIPVLHEM 60
           ||||||||||:|| |:|||| ||| |||||||||||||||||| |||||| | | ||||
Db       1 CDLPQTHSLGNRRALILLAQMGRISHFSCLKDRHDFRFPEEEFDGHQFQKTQAISVLHEM 60

Qy      61 IQQTFNLFSTKDSSAAWDETLLDKFYTELYQQLNDLEACVIQGVGVTETPLMKEDSILAV 120
           ||| |||||:|||||:::|:|:| |||||||||||||||| ||| |||||  ||||||
Db      61 IQQTFNIFSTEDSSAAWIQSLLPKFSTELYEQLNDLEACVIQEVGVEETPLMNVDSILAV 120

Qy     121 RKYFQRITLYLTEKKYSPCAWEVVRAEIMRSLSFSTNLQKRLRRKD 166
           |||||||||| |||||||||||||||||| | ||||| || |:
Db     121 RKYFQRITLYLTEKKYSPCAWEVVRAEIMRSLSFSTNLQKRLRRKD 166


RESULT 15
US-09-977-034-10
; Sequence 10, Application US/09977034
; Patent No. US20020081664A1
; GENERAL INFORMATION:
;  APPLICANT: Lo, Kin-Ming
;  APPLICANT:  Sun, Yaping
;  APPLICANT:  Gillies, Stephen D.
;  TITLE OF INVENTION: Expression and Export of Interferon-Alpha Proteins as
;  TITLE OF INVENTION:  Fc Fusion Proteins
;  FILE REFERENCE: LEX-009
;  CURRENT APPLICATION NUMBER: US/09/977,034
;  CURRENT FILING DATE:  2001-10-11
;  PRIOR APPLICATION NUMBER: US/09/575,583
;  PRIOR FILING DATE: 2000-05-19
;  PRIOR APPLICATION NUMBER: US 60/134,895
;  PRIOR FILING DATE: 1999-05-19
;  NUMBER OF SEQ ID NOS: 29
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 29
;   LENGTH: 166
;   TYPE: PRT
```



OM protein - protein search, using sw model

Run on: January 17, 2003, 06:50:48 ; Search time [illegible] Seconds
(without alignments)
692.560 Million cell updates/sec

Title: US-09-915-873-4
Perfect score: 960
Sequence: 1 MALTFALLVALLVLSCKSSC..........[illegible] 188

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

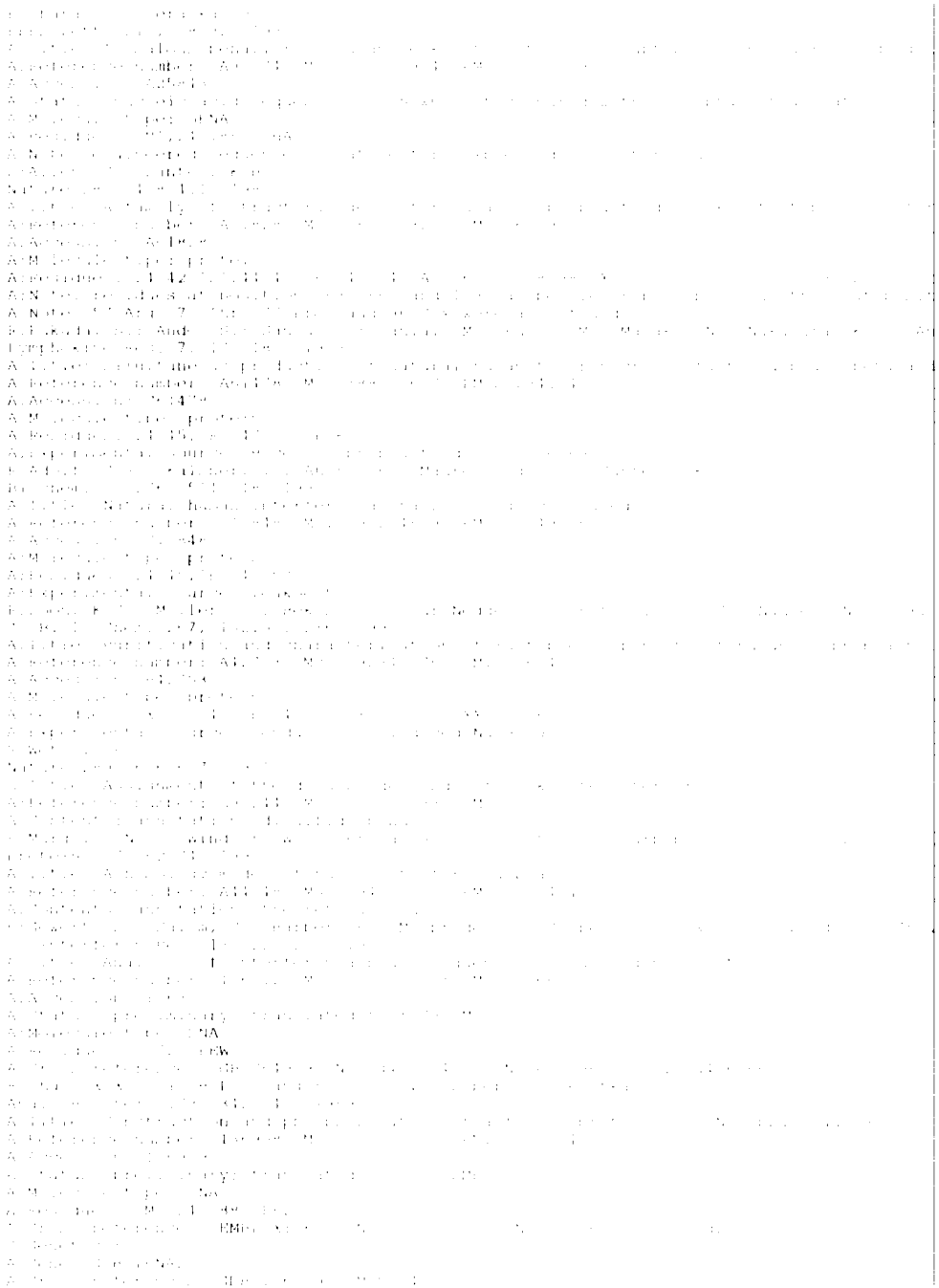
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	957	99.7	188	1	IVHUA2	interferon alpha 2
2	851	88.6	165	2	I78570	alpha 2 interferon
3	809.5	84.3	189	1	IVHUA7	interferon alpha-5
4	808.5	84.2	189	1	IVHU16	interferon alpha [illegible]
5	793.5	82.7	189	1	IVHU14	interferon alpha [illegible]
6	781.5	81.4	189	2	I52347	interferon alpha M
7	779.5	81.2	189	2	I51970	interferon precurs
8	772.5	80.5	189	1	IVHUA1	interferon alpha-1
9	770.5	80.3	189	1	IVHU4B	interferon alpha [illegible]
10	768.5	80.1	189	2	I84464	interferon-alpha-F
11	767.5	79.9	189	1	IVHU16	interferon alpha [illegible]
12	766.5	79.8	189	1	IVHUF	interferon alpha [illegible]
13	766.5	79.8	189	1	IVHUA5	interferon alpha-[illegible]
14	760.5	79.2	189	1	IVHUA9	interferon alpha-[illegible]
15	756.5	78.8	189	2	I37584	IFN alpha N protei
16	754.5	78.6	189	1	IVHU18	interferon alpha-[illegible]
17	748.5	78.0	189	2	I53102	interferon alpha [illegible]
18	744.5	77.6	189	1	IVHUA0	interferon alpha-7
19	739.5	77.0	181	2	I56313	interferon alpha 2
20	732.5	76.3	189	1	IVHUA4	interferon alpha [illegible]
21	728.5	75.9	176	2	I56314	interferon alpha -
22	722.5	75.3	194	1	IVHOA4	interferon alpha-[illegible]
23	719.5	74.9	167	2	D25843	interferon alpha-C
24	718.5	74.8	184	1	IVHOA2	interferon alpha [illegible]
25	710.5	74.0	184	1	IVHOA3	interferon alpha-[illegible]
26	704.5	73.4	184	1	IVHOA1	interferon alpha [illegible]
27	690.5	71.9	167	2	E25843	interferon alpha E
28	678.5	70.7	167	2	F25843	interferon alpha F
29	665.5	69.3	162	2	C25843	interferon alpha [illegible]
30	[illegible]	[illegible]	[illegible]	2	[illegible]	interferon alpha [illegible]
31	[illegible]	[illegible]	[illegible]	[illegible]	IVMSA[illegible]	interferon alpha [illegible]
32	[illegible]	[illegible]	[illegible]	1	IV[illegible]	interferon alpha [illegible]
33	591.5	[illegible]	[illegible]	1	IV[illegible]	interferon alpha [illegible]
34	[illegible]	[illegible]	[illegible]	[illegible]	IV[illegible]	interferon alpha [illegible]
35	[illegible]	[illegible]	[illegible]	[illegible]	IV[illegible]A	interferon alpha [illegible]
36	588.5	[illegible]	189	1	IVMSA[illegible]	interferon alpha [illegible]
37	5[illegible]	[illegible]	[illegible]	1	IV[illegible]	interferon alpha [illegible]
38	58[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	alpha interferon [illegible]
39	[illegible]	[illegible].4	[illegible]	2	A[illegible]	interferon alpha [illegible]
40	572.5	59.[illegible]	[illegible]	1	IVMSA[illegible]	interferon alpha [illegible]
41	568.5	[illegible]	192	1	IVFIA[illegible]	interferon alpha [illegible]
42	561.5	58.5	190	[illegible]	[illegible]	interferon alpha [illegible]
43	561.5	58.5	[illegible]	2	[illegible]	interferon alpha [illegible]
44	554.[illegible]	57.8	[illegible]	1	[illegible]VMSA[illegible]	interferon alpha [illegible]
45	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

ALIGNMENTS

RESULT 1
IVHUA2
interferon alpha 2 precursor [illegible]
N;Alternate names: IFN-alpha2; interferon alpha [illegible]
C;Species: Homo sapiens (man)
C;Date: 31-Oct-1985 [illegible]
C;Accession: A93234; [illegible]
R;Goeddel, D.V.; Yelverton, E.; [illegible]
ss, M.; Familletti, P.C.; Pestka [illegible]
Nature 287, 411-416, 1980
A;Title: Human leukocyte interferon [illegible]
A;Reference number: A93234; [illegible]
A;Accession: A93234
A;Molecule type: DNA
A;Residues: 1-188 [illegible]
A;Cross-references: GB:[illegible]
A;Experimental source: [illegible]
R;Goeddel, D.V.; Leung, D.W.; [illegible]
Nature 290, 20-26, 1981
A;Title: The structure of [illegible] human leukocyte interferon [illegible]
A;Reference number: A93249; [illegible]
A;Accession: [illegible]
A;Molecule type: mRNA
A;Residues: 1-188 [illegible]
A;Cross-references: [illegible]
A;Note: eight classes of interferon alpha clones were identified; [illegible]
R;Lawn, R.M.; [illegible]
Proc. Natl. Acad. Sci. U.S.A. 78, [illegible], 1981
A;Title: DNA sequence of a major human leukocyte interferon gene
A;Reference number: [illegible]
A;Accession: [illegible]
A;Molecule type: DNA
A;Residues: [illegible]
A;Cross-references: [illegible]
A;Experimental source: [illegible]
[illegible]
[illegible]
A;Title: [illegible] human leukocyte interferon [illegible]
A;Reference number: [illegible]
A;Accession: [illegible]
A;Status: preliminary; [illegible]
A;Molecule type: mRNA
A;Residues: [illegible]
A;Cross-references: [illegible]
R;Streuli, M.; Nagata, S.; Weissmann, C.
Science 209, 1343-1347, 1980
A;Title: At least three human type alpha interferons: structure [illegible]
A;Reference number: [illegible]
A;Accession: [illegible]
A;Molecule type: mRNA
A;Residues: [illegible]
A;Cross-references: [illegible]



J. Mol. Biol. 185, 227-260, 1985
A;Title: Structural relationship of human interferon alpha genes and pseudogenes.
A;Reference number: A92916; MUID:86037205; PMID:4057246
A;Accession: S43716
A;Molecule type: DNA
A;Residues: 1-189 <HEN>
A;Cross-references: EMBL:X02956; NID:g32665; PIDN:CAA26743.1; PID:g75207
R;Goeddel, D.V.; Leung, D.W.; Dull, T.J.; Gross, M.; Lawn, R.M.; McCandliss, R.; Seeb
Nature 290, 20-26, 1981
A;Title: The structure of eight distinct cloned human leukocyte interferon cDNAs.
A;Reference number: A93249; MUID:81114795; PMID:6163093
A;Accession: A01833
A;Molecule type: mRNA
A;Residues: 57-189 <GOE>
A;Cross-references: GB:V00541; GB:J00213; NID:g42718; PIDN:CAA24802.1; PID:g42719
A;Note: eight classes of interferon alpha clones were identified; this sequence is be
C;Genetics:
A;Gene: GDB:IFNA5
A;Cross-references: GDB:136862; OMIM:147565
A;Map position: 9p22-9p22
C;Superfamily: interferon alpha
C;Keywords: antiviral; cytokine; leukocyte
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-189/Product: interferon alpha 5 #status predicted <MAT>

Query Match 84.4%; Score 809.5; DB 1; Length 189;
Best Local Similarity 83.6%; Pred. No. 2.1e-67;
Matches 158; Conservative 11; Mismatches 19; Indels 1; Gaps 1;

Qy 1 MALTFALLVALLVLSCKSSCSVGCDLPQTHSLGSRRTLMLLAQMRRISLFSCLKDRHDFGFP 60
||| | ||:||:||:||| ||:|||||||||| :|||||:::|| :|: |:|:||||||
Db 1 MALPFVLLMALVVLNCKSSCSLGCDLPQTHSLSNRRTLMLMAQMSRISPFSCLKDRHDFGFP 60

Qy 61 FPQEEF-GNQFQKAETIPVLHEMIQQIFNLFSTKDSSAAWDETLLDKFYTELYQQLNDLE 119
|||||| ||||||: | |||||||| |||||||||| ||||||||||||||||||
Db 61 FPQEEFDGNQFQKAQAISVLHEMIQQTFNLFSTKDSSAIWDETLLDKFYIELYQQLNDLE 120

Qy 120 ACVIQGVGVTETPLMKEDSILAVRKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTN 179
||:| ||| :|||| |||| |||||||||||| ||||||||||||||||||| |
Db 121 ACMMQEVGVEDTPLMNVDSILVVKKYFQRITLYLTEKKYSPCAWEVVRAEIMRSFSLSAN 180

Qy 180 LQESLRSKE 188
||| || ||
Db 181 LQERLRRKE 189

RESULT 4
1VHU16
interferon alpha-1-6 precursor - human
N;Alternate names: HuIFN alpha 1-6, LeIF K, type I interferon
C;Species: Homo sapiens (man)
C;Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 18-Jun-1999
C;Accession: A23753
R;Henco, K.; Brosius, J.; Fujisawa, A.; Fujisawa, J.-I.; Haynes, J.R.; Hershberg, R.J.; K
J. Mol. Biol. 185, 227-260, 1985
A;Title: Structural relationship of human interferon alpha genes and pseudogenes.
A;Reference number: A92916; MUID:86037205; PMID:4057246
A;Accession: A23753
A;Molecule type: DNA
A;Residues: 1-189 <HEN>
A;Cross-references: GB:X02958; NID:g32662; PIDN:CAA26744.1; PID:g75204
C;Genetics:
A;Gene: GDB:IFNA6
A;Cross-references: GDB:136863; OMIM:147566
A;Map position: 9p22-9p22
C;Superfamily: interferon alpha
C;Keywords: antiviral
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-189/Product: interferon alpha 1 6 #status predicted <MAT>
F;24-122,52 162/Disulfide bonds: #status predicted

Query Match 84.2%; Score 808.5; DB 1; Length 189;
Best Local Similarity 85.7%; Pred. No. 2.6e-67;

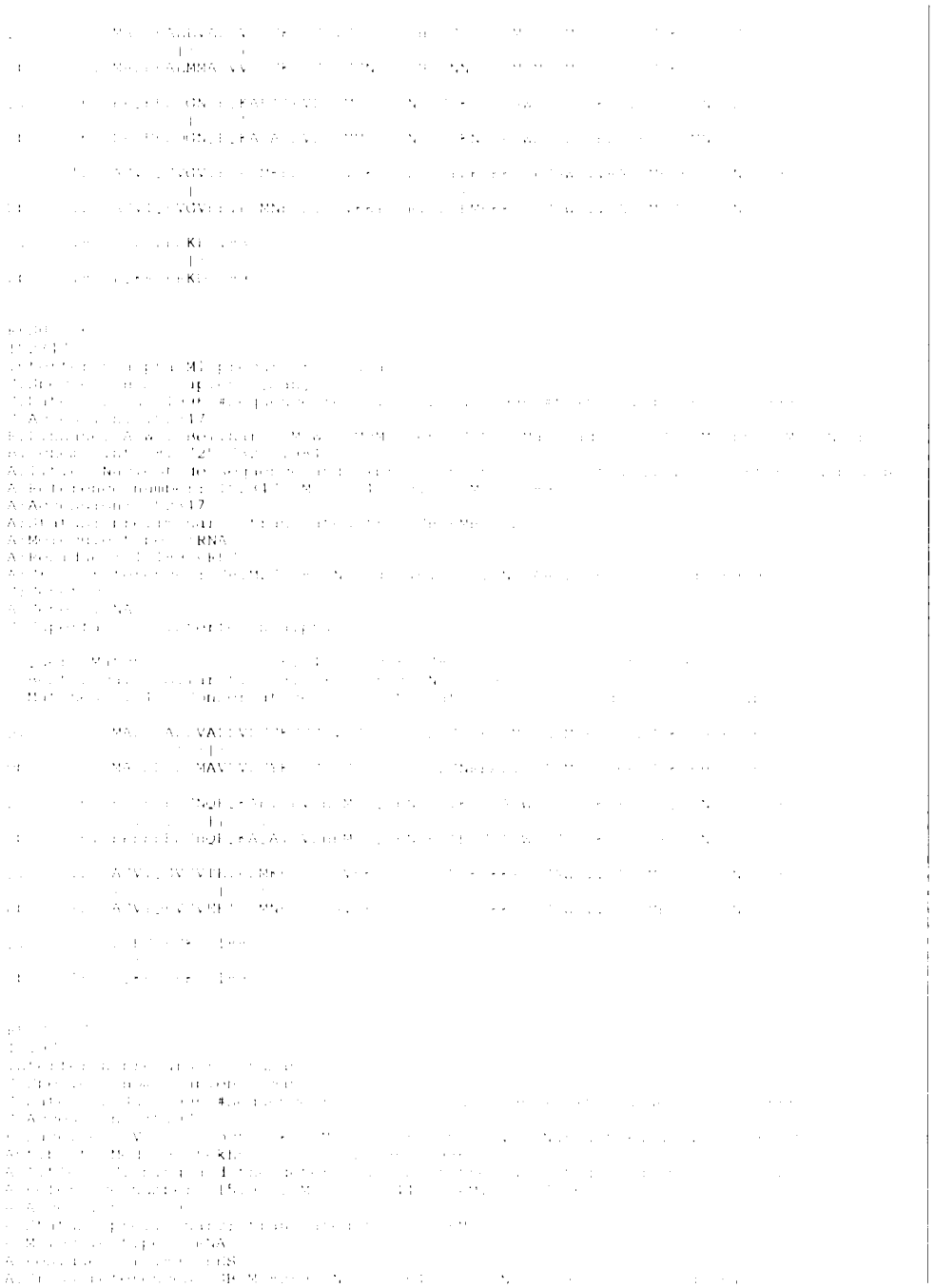
Matches [illegible]

Qy 1 MALTFALLVALLVLSCKSSCSVGCDLPQTHSLGSRRTLMLLAQMRRISLFSCLKDRHDFGFP 60
||| | ||:||:||:||| ||:|||||||||| :|||||:::|| :|: |:|:||||||
Db 1 MALPFALLMALVVLSCKSSCSLGCDLPQTHSLSNRRTLMLMAQMSRISPFSCLKDRHDFGFP 60

Qy [illegible]
[illegible]
Db [illegible]

RESULT 5
1VHU14
interferon alpha 1-14 precursor [validated] - human
N;Alternate names: HuIFN alpha 1 [illegible]; LeIF type I interferon
C;Species: Homo sapiens (man)
C;Date: [illegible]-Sep 1991 #sequence_revision [illegible] #text_change [illegible]
C;Accession: A23755; A94155; B94155; [illegible]; A97963; [illegible]
R;Henco, K.; Brosius, J.; Fujisawa, A.; Fujisawa, J.-I.; Haynes, J.R.; [illegible]
J. Mol. Biol. 185, 227-260, 1985
A;Title: Structural relationship of human interferon alpha genes and pseudogenes.
A;Reference number: A92916; MUID:86037205; PMID:4057246
A;Accession: A23755
A;Molecule type: DNA
A;Residues: 1-189 <HEN>
A;Cross-references: GB:X02957; NID:g32663; PIDN:CAA26745.1; PID:g75205
R;Goeddel, D.V.; Leung, D.W.; Dull, T.J.; Gross, M.; Lawn, R.M.; McCandliss, R.; Seeb
Nature 290, 20-26, 1981
A;Title: The structure of eight distinct cloned human leukocyte interferon cDNAs.
A;Reference number: A93249; MUID:81114795; PMID:6163093
A;Accession: B93249
A;Molecule type: mRNA
A;Residues: 1-174,'T',176-189 <LAW>
A;Cross-references: GB:V00544; GB:J00214; NID:g42720; PIDN:CAA24805.1; PID:g42721
A;Note: a variant sequence differs in [illegible] at [illegible]; [illegible]
R;Shirono, H.; Kono, J.; Omata, H.; Matsui, A.
Biosci. Biotechnol. Biochem. 56, [illegible]
A;Title: Identification of glycosylated subtypes of interferon alpha [illegible]
A;Reference number: P22203; MUID:[illegible]; PMID:[illegible]
A;Accession: P22203
A;Molecule type: protein
A;Residues: 'X',25-[illegible]
A;Experimental source: [illegible]
C;Genetics:
A;Gene: GDB:IFNA14
A;Cross-references: GDB:[illegible]; OMIM:[illegible]
A;Map position: 9p22-9p22
C;Superfamily: interferon alpha
C;Keywords: antiviral; [illegible]
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-189/Product: interferon alpha [illegible] #status predicted <MAT>
F;24-122,52 162/Disulfide bonds: #status predicted
F;25/CHO binding site: carbohydrate (Asn, covalent) #status predicted

Query Match [illegible]
Best Local Similarity [illegible]
Matches [illegible]



J. Mol. Biol. 185, 227-260, 1985
A;Title: Structural relationship of human interferon alpha genes and pseudogenes.
A;Reference number: A92916; MUID:86077205; PMID:3957246
A;Accession: S43715
A;Molecule type: DNA
A;Residues: 1-189 <HEN>
A;Cross-references: EMBL:X75934
R;Rostoks, N.
submitted to the EMBL Data Library, December 1993
A;Reference number: S41196
A;Accession: S41196
A;Molecule type: DNA
A;Residues: 1-9,'A',11-189 <ROS>
A;Cross-references: EMBL:X75934; NID:g439666; PIDN:CAA53549.1; PID:g439667
C;Genetics:
A;Gene: GDB:IFNA1
A;Cross-references: GDB:146353; OMIM:147660
A;Map position: 9p22-9p22
C;Superfamily: interferon alpha
C;Keywords: antiviral; cytokine; leukocyte
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-189/Product: interferon alpha-1 #status predicted <MAT>
F;24-122,52-162/Disulfide bonds: #status predicted

Query Match 80.5%; Score 772.5; DB 1; Length 189;
Best Local Similarity 82.0%; Pred. No. 5.4e-64;
Matches 155; Conservative 10; Mismatches 23; Indels 1; Gaps 1;

Qy 1 MALTFALVALLVLSCKSSCSVGCDLPQTHSLGSRRTLMLLAQMRRISPFSCLKDRHDFG 60
Db 1 MASPFALLMVLVVLSCKSSCSLGCDLPETHSLDNRRTLMLLAQMSRISPSSCLMDRHDFG 60

Qy 61 FPQEEFDGNQFQKAPAISVLHELIQQIFNLFTTKDSSAAWDEDLLDKFCTELYQQLNDLE 120
Db 61 FPQEEFDGNQFQKAPAISVLHELIQQIFNLFTTKDSSAAWDEDLLDKFCTELYQQLNDLE 120

Qy 120 ACVIQGVGVTETPLMKEDSILAVRKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTN 179
Db 121 ACVMQEEPVGETPLMNADSILAVKKYFRRITLYLTEKKYSPCAWEVVRAEIMRSLSLSTN 180

Qy 180 LQESLRSKE 188
Db 181 LQERLRRKE 189

RESULT 9
IVHU4B
interferon alpha-1-4b precursor - human
N;Alternate names: HuIFN-alpha-1-4b; type I interferon
C;Species: Homo sapiens (man)
C;Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 16-Jun-1994
C;Accession: E23753
R;Henco, K.; Brosius, J.; Fujisawa, A.; Fujisawa, J.I.; Haynes, J.R.; Hochstadt, J.; K
J. Mol. Biol. 185, 227-260, 1985
A;Title: Structural relationship of human interferon alpha genes and pseudogenes.
A;Reference number: A92916; MUID:86077205; PMID:3957246
A;Accession: E23753
A;Molecule type: DNA
A;Residues: 1-189 <HEN>
A;Cross-references: GB:X02955; NID:g42656; PIDN:CAA26701.1; PID:g27867
C;Genetics:
A;Gene: GDB:IFN1
A;Cross-references: GDB:119328; OMIM:147660
A;Map position: 9p22-9p22
C;Superfamily: interferon alpha
C;Keywords: antiviral
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-189/Product: interferon alpha-1-4b #status predicted <MAT>
F;24-122,52-162/Disulfide bonds: #status predicted

Query Match 80.3%; Score 770.5; DB 1; Length 189;
Best Local Similarity 80.4%; Pred. No. 8.2e-64;
Matches 152; Conservative 17; Mismatches 19; Indels 1; Gaps 1;

Qy 1 MALTFALVALLVLSCKSSCSVGCDLPQTHSLGSRRTLMLLAQMRRISPFSCLKDRHDFG 60
Db 1 MALSFSLLMAVLVLSYKSICSLGCDLPETHSLDNRRTLVLLGQMGRISPLSCLKDRKDFR 60

Qy 61 FPQEEFDGNQFQKAEIPVLHELIQQIFNLFSTKDSSAAWDETLLDKFYTELYQQLNDLE 120
Db 61 FPQEEFDGNQFQKAQAISVLHEMIQQTFNLFSTKDSSAAWDQSLLEKFSTELYQQLNDLE 120

Qy 120 ACVIQGVGVTETPLMKEDSILAVRKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTN 179
Db 121 ACVIQEVGVVETPLMNVDSILAVKKYFRRITLYLTEKKYSPCAWEVVRAEIMRSLSLSTN 180

Qy 180 LQESLRSKE 188
Db 181 LQKKLEKKE 189

RESULT 10
IFNH4
interferon alpha H - human
C;Species: Homo sapiens (man)
C;Date: 12-Apr-1995 #sequence_revision 12-Apr-1995 #text_change 16-Jun-1995
C;Accession: I81464, I37503
R;Chroni, E.Y.; Berglin, V.M.; Ishiguro, A.A.; Apostolou, H.K.; Visvardis, E.E.; Tsia
A.; Loznar, V.F.; Karpasas, V.M.; Pit..., J.; V.A.; Spandidos, D.A.
Mokl. Biochem. 269, 91-95, 1993
A;Title: A new type of human interferon alpha
A;Reference number: I37503
A;Accession: I81464
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-189 <FES>
A;Cross-references: GB:M12759; NID:g190790; PIDN:AAA52710.1; PID:g190791
A;Accession: I37503
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-189 <KRE2>
A;Cross-references: GB:...; NID:...; PIDN:AAA...; PID:...
C;Genetics:
A;Gene: IFNA
C;Superfamily: interferon alpha

Query Match ...; Score ...; DB ...; Length ...;
Best Local Similarity ...; Pred. No. ...;
Matches ...; Conservative ...; Mismatches ...; Indels ...; Gaps ...;

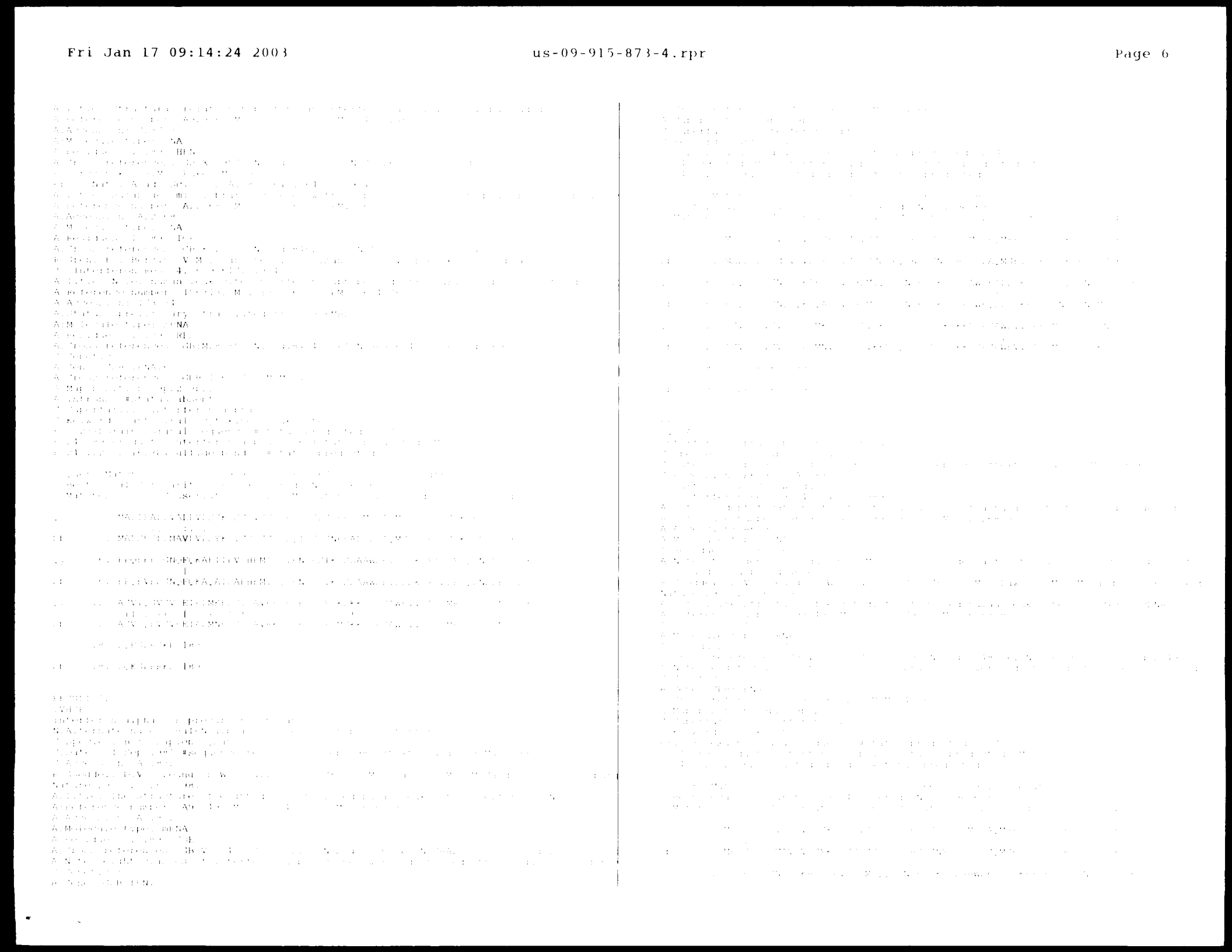
Qy 1 MALTFALVALLVLSCKSSCSVGCDLPQTHSLGSRRTLMLLAQMRRISPFSCLKDRHDFG 60
Db 1 MALSFSLLMAVLVLSYKSICSLGCDLPETHSLDNRRTLVLLAQMGRISPLSCLKDRKDFR 60

Qy 61 FPQEEFDGNQFQKAEIPVLHEMIQQIFNLFSTKDSSAAWDETLLDKFYTELYQQLNDLE 120
Db 61 FPQEEFDGNQFQKAQAISVLHEMIQQTFNLFSTKDSSAAWDQSLLEKFSTELYQQLNDLE 120

Qy 120 ACVIQGVGVTETPLMKEDSILAVRKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTN 179
Db 121 ACVIQEVGVVETPLMNVDSILAVKKYFRRITLYLTEKKYSPCAWEVVRAEIMRSLSLSTN 180

Qy 180 LQESLRSKE 188
Db 181 LQKKLEKKE 189

RESULT 11
IVHU1B
interferon alpha-1-4b precursor - human
N;Alternate names: HuIFN-alpha-1-4b; type I interferon
C;Species: Homo sapiens (man)
C;Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 16-Jun-1994
C;Accession: E23753, A22753, I37...
R;Henco, K.; Brosius, J.; Fujisawa, A.; Fujisawa, J.I.; Haynes, J.R.; Hochstadt, J.; K
J. Mol. Biol. 185, 227-260, 1985



```
Db      61 IPQEEFDGNQFQKAQAISVLHEMIQQTFNLFSTKDSSAAWDETLLDKFYIELYQQLNDLE 120

Qy     120 ACVIQGVGVTETPLMKEDSILAVRKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTN 179
           ||||| ||| ||||| |||||||||||||||||| |:|||||||||||.|||| | |||
Db     121 ACVIQEVGVEETPLMNEDSILAVRKYFQRITLYLIERKYSPCAWEVVRAEIMRSLSFSTN 180

Qy     180 LQESLRSKE 188
           ||: || |:
Db     181 LQKRLRRKD 189


RESULT 14
IVHUA9
interferon alpha-17 precursor - human
N;Alternate names: interferon alpha 9; interferon alpha-I'
C;Species: Homo sapiens (man)
C;Date: 01-Sep-1981 #sequence_revision 01-Sep-1981 #text_change 21-Jul-2000
C;Accession: A01835; A22255; C42753
R;Lawn, R.M.; Adelman, J.; Bell, L.J.; Gross, M.; Goeddel, D.; Ullrich, A.
Science 212, 1159-1162, 1981
A;Title: DNA sequence of two closely linked human leukocyte interferon genes.
A;Reference number: A94255; MUID:81201124; PMID:6165002
A;Accession: A01835
A;Molecule type: DNA
A;Residues: 1-189 <LAW>
A;Cross-references: GB:J00214; GB:V00532; NID:g22622; PIDN:AAA52797.1; PID:g28474
R;Mizoguchi, J.; Pitha, P.M.; Raj, N.B.K.
DNA 4, 221-232, 1985
A;Title: Efficient expression in Escherichia coli of two species of human interferon-alp
A;Reference number: A22255; MUID:85220965; PMID:2991575
A;Accession: A22255
A;Molecule type: mRNA
A;Residues: 1-56,'H',58-189 <MIZ>
A;Cross-references: GB:M11028; NID:g184612; PIDN:AAA52725.1; PID:g306903
R;Zoon, K.C.; Miller, E.; Bekisz, J.; zur Nedden, D.; Rubinstein, S.; Familletti, P.C.
J. Biol. Chem. 267, 15210-15216, 1992
A;Title: Purification and characterization of multiple components of human lymphoblast
A;Reference number: A42753; MUID:92340576; PMID:1634550
A;Accession: C42753
A;Molecule type: protein
A;Residues: 'X',25-50,'XX',53-56 <ZOO>
C;Genetics:
A;Gene: GDB:IFNA17
A;Cross-reference: GDB:136358; OMIM:147583
A;Map position: 9p22-9p22
C;Superfamily: interferon alpha
C;Keywords: leukocyte
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-189/product: interferon alpha-17 #status predicted <MAT>
F;21 122,52-162/Disulfide bonds: #status predicted

  Query Match             79.2%;  Score 760.5;  DB 1;  Length 189;
  Best Local Similarity   79.9%;  Pred. No. 6.9e-63;
  Matches  151;  Conservative   17;  Mismatches   20;  Indels    1;  Gaps    1;

Qy       1 MALIFALLVALLVLSCKSSCSVGCDLPQTHSLGSRRTLMLLAQMRRISLFSCLKDRHDFG 60
           |||:|||:|:|||| || ||:|||||||||||||:|| :|||||  | | || | | : | |
Db       1 MALSFSLLMAVLVLSYKSICSLGCDLPQTHSLGNRRALILLAQMGRISPFSCLKDRKDFG 60

Qy      61 FPQEEF-GNQFQKAETIPVLHEMIQQTFNLFSTKDSSAAWDETLLDKFYIELYQQLNDLE 119
            ||||| |||||| : | |||||||| |||||||:|||||:::|':|| ||||:|||: |
Db      61 LPQEEFDGNQFQKIQAISVLHEMIQQTFNLFSTEDSSAAWEQSLLEKFSIELYQQLNNLE 120

Qy     120 ACVIQGVGVTETPLMKEDSILAVRKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTN 179
           ||||| || ||||| |||||||||||||||||| |||||||| |||:|||| |  |||
Db     121 ACVIQEVGMEETPLMNEDSILAVRKYFQRITLYLIEKKYSPCAWEVVRAEIMRSLSFSTN 180

Qy     180 LQESLRSKE 188
           ||: || |:
Db     181 LQKILRRKD 189


RESULT 15
```

```
; MF_01
IFN-alpha-N precursor - human
C;Species: Homo sapiens (man)
C;Date: 04-Apr-1995 #sequence_revision [illegible] #text_change [illegible]
C;Accession: I37684
R;Rubin, B.Y.; Borecky, V.M.; Ishenko, A.V.; Apostolov, O.K.; Vinogradova, [illegible]
A.A.; Lozko, V.P.; Karasev, V.M.; Ilin, [illegible] V.A.; Sverdlov, E.D.
DokI. Biochem. 269, 91-95, 1983
A;Title: A new type of leukocyte interferon.
A;Reference number: I37684
A;Accession: I37684
A;Status: preliminary; translated [illegible] 30-May-1995
A;Molecule type: mRNA
A;Residues: 1-189 <RUB>
A;Cross-references: EMBL:X01138; NID:g31; PIDN:AAA51[illegible]; [illegible]
C;Superfamily: interferon alpha

  Query Match             [illegible]%;  Score [illegible];  DB [illegible];  Length [illegible];
  Best Local Similarity   [illegible]%;  Pred. No. [illegible];
  Matches  [illegible];  Conservative  [illegible];  Mismatches  [illegible];  Indels  [illegible];  Gaps  [illegible];

Qy       1 MALIFALLVALLVLSCKSSCSVGCDLPQTHSLGSRRTLMLLAQMRRISLFSCLKDRHDFG 60
           | |:|:|| |:|||  |  |:|  |  |  |  |  |  |  |  |  |  |  | |
Db       1 MPLSFSLLMAVLVLSYKSICSLGCDLPQTHNLRNRRALILLAQMGRISPFSCLKDRKDFR 60

Qy      61 FPQEEF-GNQFQKAETIPVLHEMIQQTFNLFSTKDSSAAWDETLLDKFYIELYQQLNDLE 119
           | | : |:|: | :  |  |  |  |  |  |  |  |  |  |  |  |  |  | :
Db      61 FPQEVFDGNQFQKAQAISVLHEMIQQIFNLFSTKDSSAAWDQTLLDKFYTELYQQLNDLE 120

Qy     120 ACVIQGVGVTETPLMKEDSILAVRKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTN 179
           | | | |:|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db     121 ACVIQEVGVEETPLMNEDSILAVRKYFQRITLYLIEKKYSPCAWEVVRAEIMRSLSFSTN 180

Qy     180 LQESLRSKE 188
           ||: || |:
Db     181 LQKILRRKD 189


Search completed: [illegible]
Job time : [illegible] secs
```

OM protein - protein search, using sw model

Run on: January 17, 2003, 05:24:45 ; Search time 14.7111 Seconds
(without alignments)
585.646 Million cell updates/sec

Title: US-09-915-873-4
Perfect score: 960
Sequence: 1 MALTFALLVALLVLSCKSSC.....EIMESTQLSING.ESLLNKL 189

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476628 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

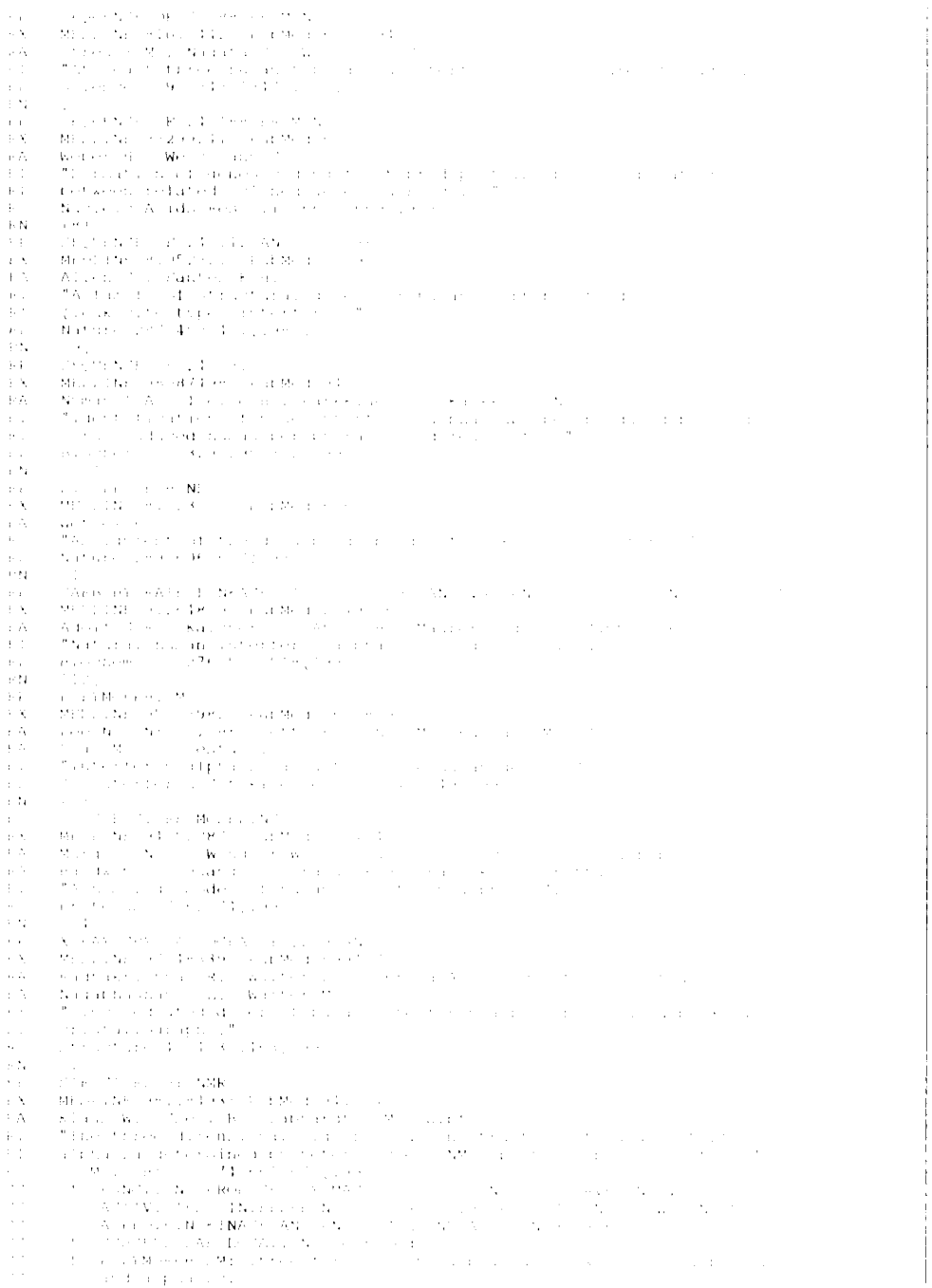
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	957	99.7	189	1	INA2_HUMAN	P05566 homo sapiens
2	899.5	84.3	189	1	INA5_HUMAN	P05568 homo sapiens
3	808.5	84.2	189	1	INA6_HUMAN	P05019 homo sapiens
4	793.5	82.7	189	1	INAD_HUMAN	P05570 homo sapiens
5	772.5	80.5	189	1	INA1_HUMAN	P05562 homo sapiens
6	770.5	80.3	189	1	INA4_HUMAN	P05014 homo sapiens
7	770.5	80.3	189	1	INA3_HUMAN	P05571 homo sapiens
8	768.5	80.1	189	1	INAK_HUMAN	P05568 homo sapiens
9	767.5	79.9	189	1	INAF_HUMAN	P05015 homo sapiens
10	766.5	79.8	189	1	INAA_HUMAN	P05566 homo sapiens
11	754.5	78.6	189	1	INAB_HUMAN	P52881 homo sapiens
12	744.5	77.5	189	1	INA7_HUMAN	P05567 homo sapiens
13	722.5	75.3	184	1	INA4_HORSE	P05006 equus caball
14	718.5	74.8	184	1	INA2_HORSE	P05004 equus caball
15	710.5	74.0	184	1	INA3_HORSE	P05005 equus caball
16	704.5	73.4	184	1	INA1_HORSE	P05003 equus caball
17	648.5	66.7	189	1	INA1_PIG	P63077 sus scrofa
18	603.5	62.9	189	1	INA5_MOUSE	P07449 mus musculus
19	596.5	62.1	189	1	INA1_BOVIN	P05563 bos taurus
20	595.5	62.0	189	1	INAB_BOVIN	P05565 bos taurus
21	591.5	61.6	189	1	INAD_BOVIN	P05571 bos taurus
22	590.5	61.5	189	1	INAA_BOVIN	P05567 bos taurus
23	590.5	61.5	189	1	INAB_BOVIN	P05568 bos taurus
24	588.5	61.3	189	1	INA1_MOUSE	P05172 mus musculus
25	582.5	60.7	189	1	INA7_BOVIN	P63077 bos taurus
26	581.5	60.6	189	1	INAC_BOVIN	P05569 bos taurus
27	581.5	60.6	189	1	INAF_BOVIN	P63076 bos taurus
28	579.5	60.4	190	1	INA9_MOUSE	P07435 mus musculus
29	572.5	59.6	190	1	INA2_MOUSE	P05173 mus musculus
30	568.5	59.2	192	1	INA1_RAT	P05121 rattus norv
31	561.5	58.5	190	1	INA7_MOUSE	P06799 mus musculus
32	554.5	57.8	189	1	INA6_MOUSE	P07450 mus musculus
33	552.5	57.6	189	1	INA8_MOUSE	P17660 mus musculus
34	518.5	54.0	189	1	INA9_HUMAN	P05565 homo sapiens
35	517.5	53.9	189	1	INAG_HUMAN	P05572 homo sapiens
36	517	53.9	191	1	INA1_CHICK	P05566 gallus gallus
37	492.5	51.3	195	1	INA_BUFMA	P05567 bufo marinus
38	472.5	49.2	197	1	INA1_BOVIN	P05564 bos taurus
39	470.5	49.0	201	1	INA1_XENLA	P05565 xenopus laevis
40	439.5	45.8	187	1	INA1_XENLA	P05563 xenopus laevis
41	427.5	44.5	198	1	INA1_HUMAN	P05564 homo sapiens
42	427.5	44.5	198	1	INA1_BOVIN	P05570 bos taurus
43	396.5	41.3	194	1	INA_TORMA	P05569 torpedo marmorata
44	389.5	40.6	194	1	INA_TORCA	P05568 torpedo californica
45	385.5	40.1	194	1	INA_TORNO	P05567 torpedo nobiliana

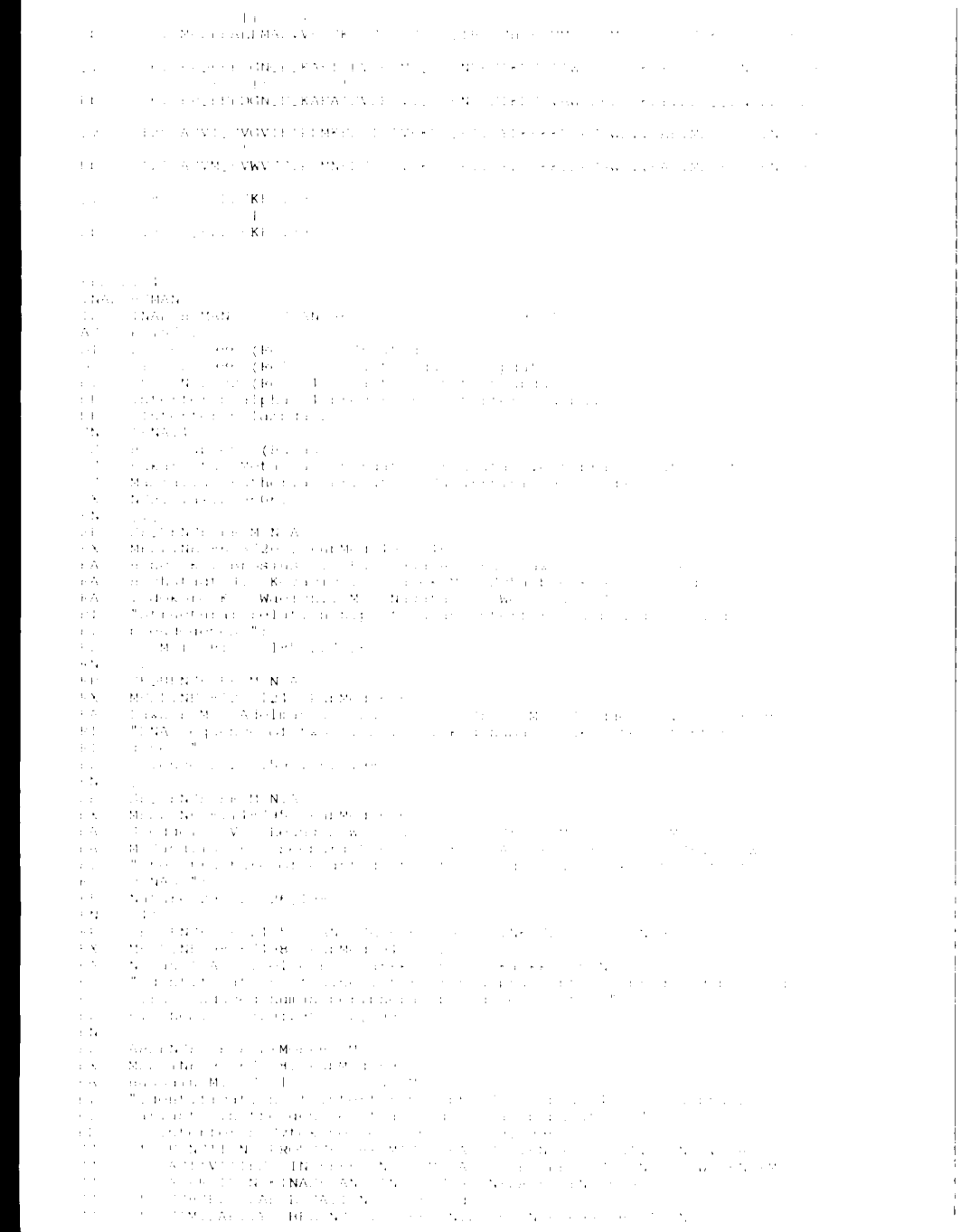
ALIGNMENTS

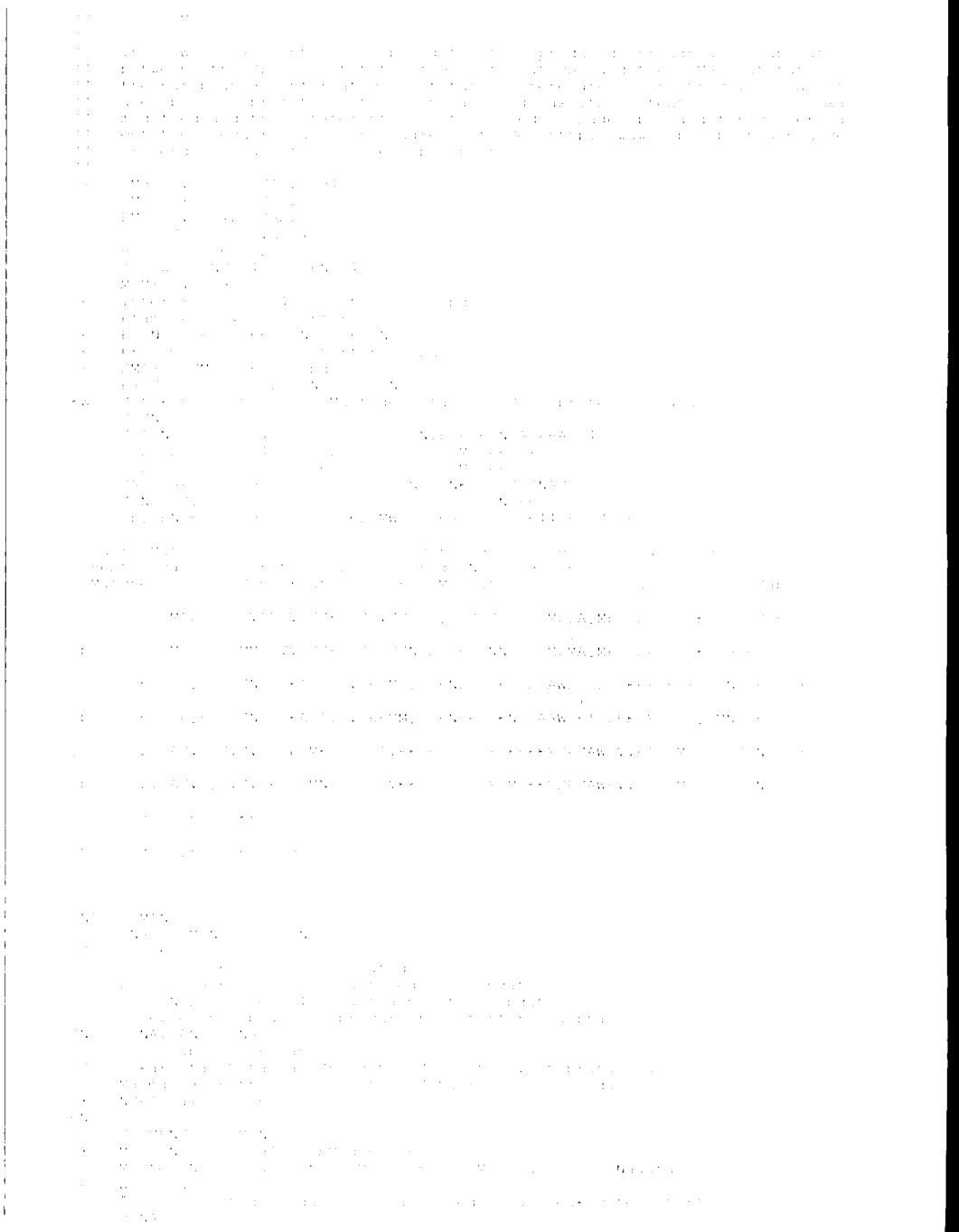
RESULT 1
INA2_HUMAN
ID INA2_HUMAN STANDARD; PRT; 189 AA.
AC P05566; Q07541; Q15 [illegible]
DT 21-JUL-1988 (Rel. 07, Created)
DT 21-JUL-1988 (Rel. 07, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Interferon alpha-2 precursor (IFN-alpha-2) (Leukocyte A) (Leu 2).
GN IFNA2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; [illegible]
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; H.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83072422; PubMed=6401927;
RA Goeddel D.V., Yelverton E., Ullrich A., Heyneker H.L., Miozzari G.,
RA Holmes W., Seeburg P.H., Dull T., May L., Stebbing N., Crea R.,
RA Maeda S., McCandliss R., Sloma A., Tabata J.M., Gross M.,
RA Familletti P.C., Pestka S.;
RT "Human leukocyte interferon produced [illegible] [illegible]
RT active.";
RL Nature 287:411-416(1980).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=84104675; PubMed=6319135;
RA Goeddel D.V., Leung D.W., Dull T.J., Gross M., Lawn R.M.,
RA McCandliss R., Seeburg P.H., Ullrich A., Yelverton E., Gray P.W.;
RT "The structure of the human fibroblast interferon gene.";
RT "DNAs.";
RL Nature 290:20-26(1981).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=82060610; PubMed=6285148;
RA Lawn R.M., Gross M., Houck C.M., Franke A.E., Gray P.W.,
RA Goeddel D.V.;
RT "DNA sequence of a human leukocyte interferon cDNA.";
RL Proc. Natl. Acad. Sci. U.S.A. 78:5435-5439(1981).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Bone marrow [illegible];
RX MEDLINE=86166451; PubMed=3456192;
RA [illegible] Barbosa [illegible] Varela [illegible] X. [illegible]
RT "Cloning of human [illegible] interferon cDNA and [illegible]
RT [illegible] in [illegible]";
RL Rev. Latinoam. Microbiol. [illegible] (1985).
RN [5]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=86205411; PubMed=3010248;
RA Austria R., Bonelli F., Cantaluppi M.L., Manzoni P., Pellegrini G.,
RA Mannoni P., Quaranta G.;
RT "A deletion [illegible] [illegible] human interferon-alpha
RT transduced human leukocytes [illegible]";
RL Cancer Gene Ther. 1:21-27(1984).
RN [6]



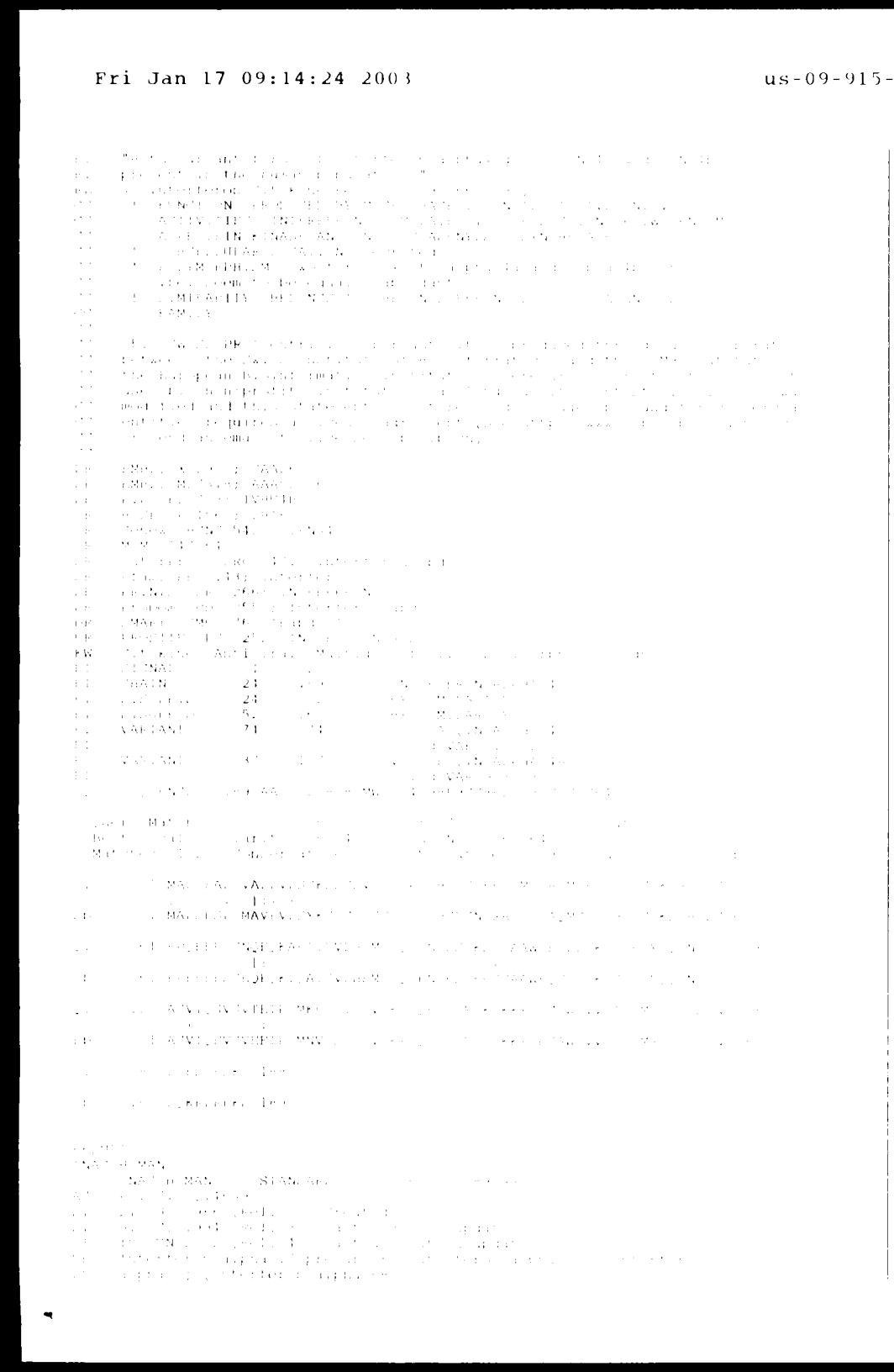
```
ID   INA5_HUMAN     STANDARD;      PRT;   189 AA.
AC   P01569;
DT   21-JUL-1986 (Rel. 01, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Interferon alpha-5 precursor (Interferon alpha G) (IeIF G) (Interferon
DE   alpha-61).
GN   IFNA5.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86037205; PubMed=4057246;
RA   Henco K., Brosius J., Fujisawa A., Fujisawa J.-I., Haynes J.R.,
RA   Hochstadt J., Kovarik T., Pasek M., Schambock A., Schmidt T.,
RA   Todokoro K., Waelchli M., Nagata S., Weissmann C.;
RT   "Structural relationship of human interferon alpha genes and
RT   pseudogenes.";
RL   J. Mol. Biol. 185:227-260(1985).
RN   [2]
RP   SEQUENCE OF 57-189 FROM N.A.
RC   TISSUE=Spleen;
RX   MEDLINE=81148795; PubMed=6163983;
RA   Goeddel D.V., Leung D.W., Dull T.J., Gross M., Lawn R.M.,
RA   McCandliss R., Seeburg P.H., Ullrich A., Yelverton E., Gray P.W.;
RT   "The structure of eight distinct cloned human leukocyte interferon
RT   cDNAs.";
RL   Nature 290:20-26(1981).
CC   -!- FUNCTION: PRODUCED BY MACROPHAGES, IFN ALPHA HAVE ANTIVIRAL
CC       ACTIVITIES, INTERFERON STIMULATES THE PRODUCTION OF TWO ENZYMES:
CC       A PROTEIN KINASE AND AN OLIGOADENYLATE SYNTHETASE.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: BELONGS TO THE INTERFERON ALPHA, BETA AND DELTA
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch)
CC   ---------------------------------------------------------------------
DR   EMBL; X02956; CAA26702.1; -.
DR   EMBL; V00541; CAA23802.1; -.
DR   PIR; A01833; IVHUA7.
DR   HSSP; P01563; 2HIE.
DR   Genew; HGNC:5426; IFNA5.
DR   MIM; 147565; -.
DR   InterPro; IPR000471; Interferon_abd.
DR   Pfam; PF00143; interferon; 1.
DR   PRINTS; PR00266; INTERFERONAB.
DR   ProDom; PD000550; Interferon_abd; 1.
DR   SMART; SM00076; IFabd; 1.
DR   PROSITE; PS00252; INTERFERON_A_B_D; 1.
KW   Cytokine; Antiviral; Multigene family; Signal.
FT   SIGNAL        1     23
FT   CHAIN        24    189       INTERFERON ALPHA 5.
FT   DISULFID     24    122       BY SIMILARITY.
FT   DISULFID     52    162       BY SIMILARITY.
SQ   SEQUENCE   189 AA;  21942 MW;  [illegible] CRC64;

  Query Match             84.4%;  Score 809.5;  DB 1;  Length 189;
  Best Local Similarity   83.6%;  Pred. No. 2e-67;
  Matches  158;  Conservative   11;  Mismatches   19;  Indels    1;  Gaps    1;

Qy       1 MALIFALLVALLVLSCKSSCSVGCDLPQTHSLGSRRTLMLLAQMRRISLFSCLKDRHDFG 60
           ||| | ||:||:|||:|| ||:|||||||||| :|||||:|:|| ||| ||||||||||
Db       1 MALPPVLLMALVVLNCKSICSLGCDLPQTHSLSNRRTLMLMAQMGRISPFSCLKDRHDFG 60

Qy      61 FPQEEFDGNQFQKAQAISVLHEMIQQTFNLFSTEHSSAAWDETLLDKFYTELYQQLNDLE 120
```

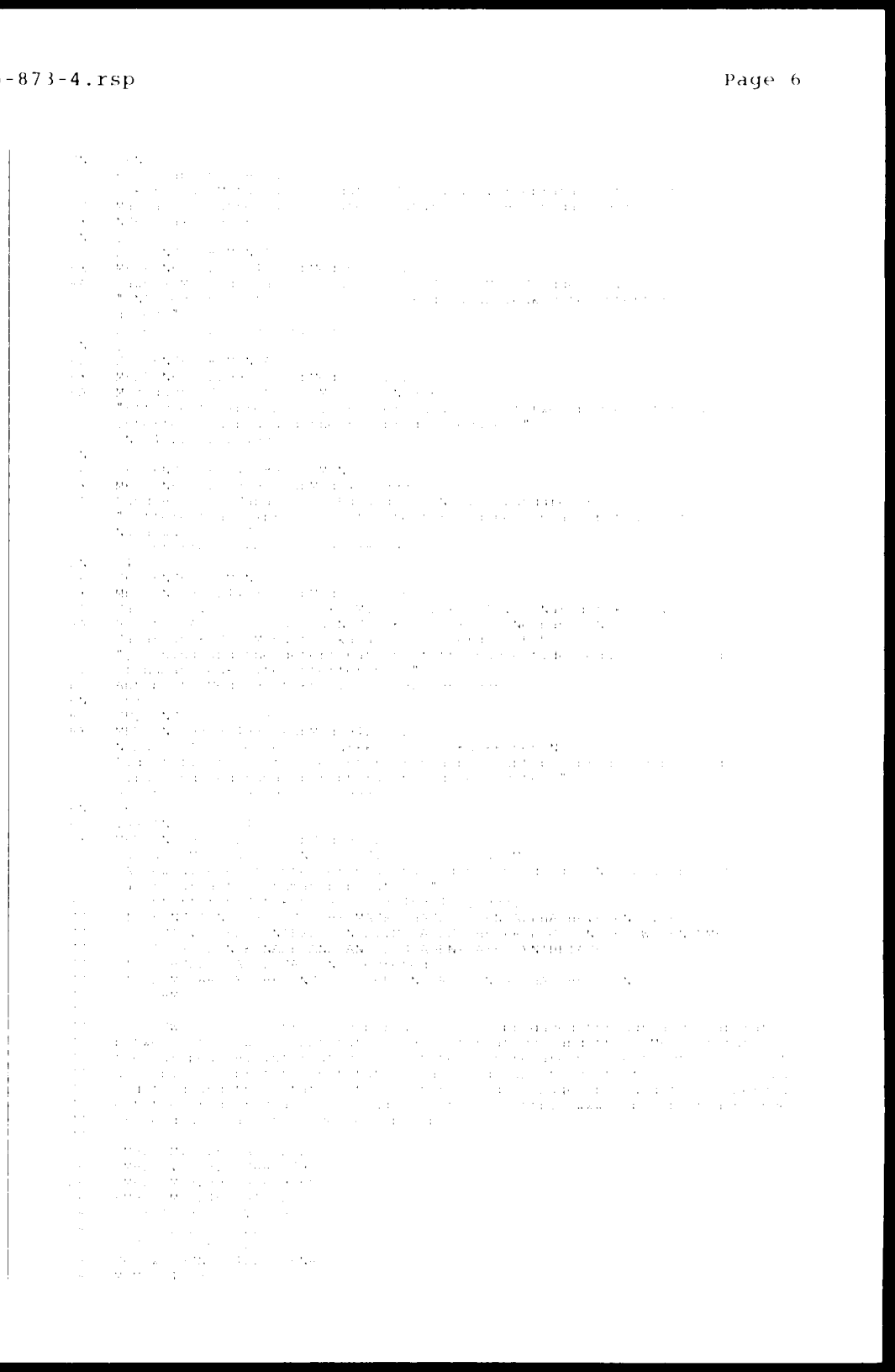
```
ID   INA6_HUMAN     STANDARD;      PRT;   [illegible] AA.
AC   P05013;
DT   [illegible] (Rel. 15, Created)
DT   15-AUG-1987 (Rel. 05, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Interferon alpha-6 precursor (Interferon alpha-6) (IeIF 6) (Interferon
DE   alpha-54).
GN   IFNA6.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86037205; PubMed=4057246;
RA   Henco K., Brosius J., Fujisawa A., Fujisawa J.-I., Haynes J.R.,
RA   Hochstadt J., Kovarik T., Pasek M., Schambock A., Schmidt T.,
RA   Todokoro K., Waelchli M., Nagata S., Weissmann C.;
RT   "Structural relationship of human interferon alpha genes and
RT   pseudogenes.";
RL   J. Mol. Biol. 185:227-260(1985).
RN   [2]
RP   SEQUENCE OF 57-189 FROM N.A.
RC   TISSUE=Spleen;
RX   MEDLINE=81148795; PubMed=6163983;
RA   Goeddel D.V., Leung D.W., Dull T.J., Gross M., Lawn R.M.,
RA   McCandliss R., Seeburg P.H., Ullrich A., Yelverton E., Gray P.W.;
RT   "The structure of eight distinct cloned human leukocyte interferon
RT   cDNAs.";
RL   Nature 290:20-26(1981).
CC   -!- FUNCTION: PRODUCED BY MACROPHAGES, IFN ALPHA HAVE ANTIVIRAL
CC       ACTIVITIES, INTERFERON STIMULATES THE PRODUCTION OF TWO ENZYMES:
CC       A PROTEIN KINASE AND AN OLIGOADENYLATE SYNTHETASE.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: BELONGS TO THE INTERFERON ALPHA, BETA AND DELTA
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X02956; CAA26704.1; -.
DR   PIR; A23754; IVHUI6.
DR   HSSP; P01563; 2HIE.
DR   Genew; HGNC:5427; IFNA6.
DR   MIM; 147566; -.
DR   InterPro; IPR000471; Interferon_abd.
DR   Pfam; PF00143; interferon; 1.
DR   PRINTS; PR00266; INTERFERONAB.
DR   ProDom; PD000550; Interferon_abd; 1.
DR   SMART; SM00076; IFabd; 1.
DR   PROSITE; PS00252; INTERFERON_A_B_D; 1.
KW   Cytokine; Antiviral; Multigene family; Signal.
FT   SIGNAL        1     23
FT   CHAIN        24    189       INTERFERON ALPHA 6.
FT   DISULFID     24    122       BY SIMILARITY.
FT   DISULFID     52    162       BY SIMILARITY.
SQ   SEQUENCE   189 AA;  [illegible] MW;  [illegible] CRC64;

  Query Match             84.4%;  Score 809.5;  DB 1;  Length 189;
  Best Local Similarity   83.6%;  Pred. No. 2e-67;
  Matches  158;  Conservative   11;  Mismatches   19;  Indels    1;  Gaps    1;

Qy       1 MALIFALLVALLVLSCKSSCSVGCDLPQTHSLGSRRTLMLLAQMRRISLFSCLKDRHDFG 60
           ||| | ||:||:|||:|| ||:|||||||||| :|||||:|:|| ||| ||||||||||
Db       1 MALPPVLLMALVVLNCKSICSLGCDLPQTHSLSNRRTLMLMAQMGRISPFSCLKDRHDFG 60

Qy      61 FPQEEFDGNQFQKAQAISVLHEMIQQTFNLFSTEHSSAAWDETLLDKFYTELYQQLNDLE 120
```





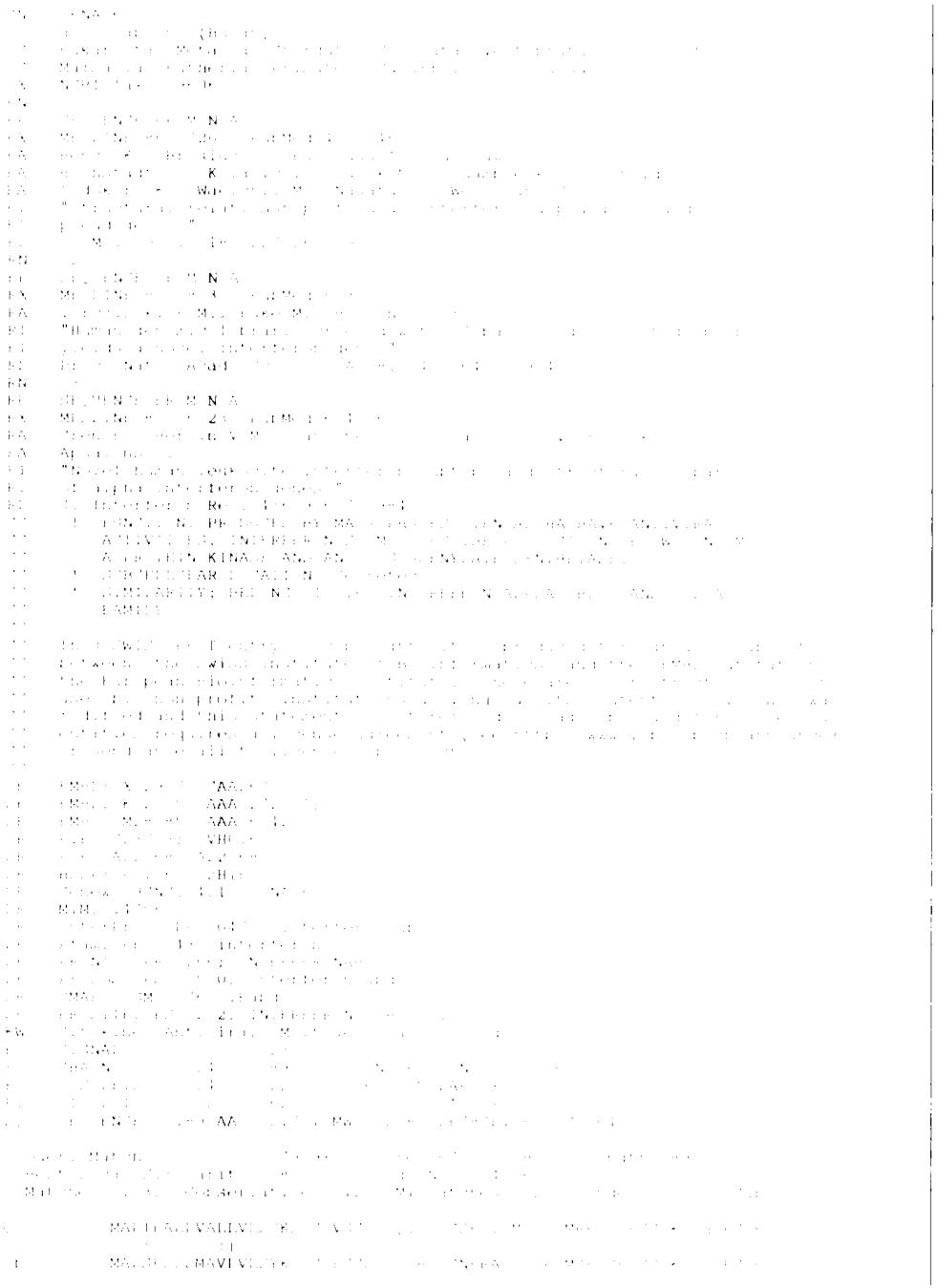
```
RT   Gene 10:1-10(1980).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=80254543; PubMed=6157095;
RA   Taniguchi T., Mantei N., Schwarzstein M., Nagata S., Muramatsu M.,
RA   Weissmann C.;
RT   "Human leukocyte and fibroblast interferons are structurally
RT   related.";
RL   Nature 285:547-549(1980).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=81148795; PubMed=6163083;
RA   Goeddel D.V., Leung D.W., Dull T.J., Gross M., Lawn R.M.,
RA   McCandliss R., Seeburg P.H., Ullrich A., Yelverton E., Gray P.W.;
RT   "The structure of eight distinct cloned human leukocyte interferon
RT   cDNAs.";
RL   Nature 290:20-26(1981).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85003592; PubMed=6479148;
RA   Todokoro K., Kioussis D., Weissmann C.;
RT   "Two non-allelic human interferon alpha genes with identical coding
RT   regions.";
RL   EMBO J. 3:1809-1812(1984).
RN   [5]
RP   SEQUENCE OF 24-58.
RX   MEDLINE=98087498; PubMed=9425112;
RA   Nyman T.A., Toppila S., Parkkinen J., Kalkkinen N.;
RT   "Identification of nine interferon-alpha subtypes produced by Sendai
RT   virus-induced human peripheral blood leucocytes.";
RL   Biochem. J. 329:295-302(1998).
RN   [6]
RP   POLYMORPHISM.
RX   MEDLINE=20485144; PubMed=11032395;
RA   Hussain M., Ni D., Gill D., Liao M.-J.;
RT   "IFN-alpha-1a gene is the major variant in the North American
RT   population.";
RL   J. Interferon Cytokine Res. 20:763-768(2000).
CC   -!- FUNCTION: PRODUCED BY MACROPHAGES. IFN ALPHA HAVE ANTIVIRAL
CC       ACTIVITIES. INTERFERON STIMULATES THE PRODUCTION OF TWO ENZYMES:
CC       A PROTEIN KINASE AND AN OLIGOADENYLATE SYNTHETASE.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- POLYMORPHISM: Two forms exist: alpha-1a (shown here) and alpha-1b.
CC   -!- MISCELLANEOUS: INTERFERONS ALPHA-1 AND ALPHA-13 HAVE IDENTICAL
CC       PROTEIN SEQUENCES.
CC   -!- SIMILARITY: BELONGS TO THE INTERFERON ALPHA, BETA AND DELTA
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J00210; AAB59403.1; -.
DR   EMBL; V00537; CAA23798.1; -.
DR   EMBL; V00538; CAA23799.1; -.
DR   EMBL; X00803; CAA25381.1; -.
DR   PIR; A01826; IVHUA1.
DR   PIR; C23285; C23285.
DR   HSSP; P01563; 2HIE.
DR   Genew; HGNC:5417; IFNA1.
DR   Genew; HGNC:5419; IFNA13.
DR   MIM; 147660; -.
DR   MIM; 147578; -.
DR   InterPro; IPR000471; Interferon_abd.
DR   Pfam; PF00143; interferon; 1.
DR   PRINTS; PR00266; INTERFERONAB.
DR   ProDom; PD000550; Interferon_abd; 1.
DR   SMART; SM00076; IFabd; 1.
DR   PROSITE; PS00252; INTERFERON_A_B_D; 1.
KW   [illegible] Antiviral; Multigene family; Signal; Polymorphism.
FT   SIGNAL       1     23
FT   CHAIN        24    189       IFN-alpha-1 ALPHA 1/13.
FT   DISULFID     24    122       BY SIMILARITY.
FT   DISULFID     92    102       BY SIMILARITY.
FT   VARIANT      149   149       A -> V (IN ALPHA 1b; dbSNP:[illegible]).
FT                                /FTId=VAR_[illegible].
SQ   SEQUENCE   189 AA;  [illegible] MW;  [illegible] CRC64;

  Query Match             [illegible]  Score [illegible]  DB 1:  [illegible]
  Best Local Similarity   [illegible]  Pred. No. [illegible]
  Matches  [illegible]  Conservative  [illegible]  Mismatches  [illegible]  Indels  [illegible]  Gaps  [illegible]

Qy       1 MALLVALLVALLVLSCKSSCSVGCDLPQTHSLGSRRTLMLLAQMRRISPFSCLKDRHDFGFPQEEFDGNQFQKAPAISVLHELIQQIFNLFTTKDSSAAWDETLLDKFYTELYQQLNDLEACVIQGVGVTETPLMKEDSILAVRKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTNLQESLRSKE
           [illegible]
Db       1 MASPFALLMVIVVLSCKSSCSLGCDLPETHSLDNRRTLMLLAQMSRISPSSCLMDRHDFGFPQEEFGNQFQKAETISVLHEVIQQTFNLFSTKDSSAAWDETLLEKFYIELFQQLNDLEACVIQEVGVEETPLMNVDSILAVKKYFQRITLYLTEKKYSPCAWEVVRAEIMRSFSLSTNLQKRLRRKD
           [illegible]
Qy      61 PVQEEFDGNQFQKAETISVLHELIQQTFNLFSTKDSSAAWDETLLEKFYIELFQQLNDLEACVIQEVGVEETPLMNVDSILAVKKYFQRITLYLTEKKYSPCAWEVVRAEIMRSFSLSTNLQKRLRRKD
Db      61 PVQEEFDGNQFQKAETISVLHELIQQTFNLFSTKDSSAAWDETLLEKFYIELFQQLNDLEACVIQEVGVEETPLMNVDSILAVKKYFQRITLYLTEKKYSPCAWEVVRAEIMRSFSLSTNLQKRLRRKD

Qy     121 AVMQEEKVDFLPLMNLDSKSRSIPQMMLNQIFLILSGSKDSTWGNTVLENLVAEVEKGDCSPVEIQENNQLEKTCKEQVTFFHDNIVTQKQSISALEHWLKLTMRSPSQNA
           [illegible]
Db     121 AVMQEEKVDFLPLMNLDSKSRSIPQMMLNQIFLILSGSKDSTWGNTVLENLVAEVEKGDCSPVEIQENNQLEKTCKEQVTFFHDNIVTQKQSISALEHWLKLTMRSPSQNA

Qy     181 LSPEDLEKLTP 189
           [illegible]
Db     181 LSPEDLEKLTP 189


RESULT 6
INA4_HUMAN
ID   INA4_HUMAN      STANDARD;      [illegible]  AA.
AC   P05014; [illegible];
DT   19-AUG-1987 (Rel. 05, Created)
DT   19-AUG-1987 (Rel. 05, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Interferon alpha-4 precursor (Interferon alpha-4B) (Interferon
DE   alpha-M1) (Interferon alpha-").
GN   IFNA4.
OS   Homo Sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86047205; PubMed=6054046;
RA   Benoit K., Brosius J., Pollack A., Ballantyne J. L., Barnes T.M.,
RA   Bachstadt J., Kacarich T., Pasek M., Schambock A., Schmitt J.,
RA   Todokoro K., Wadsehill M., Narita K., Weissmann C.;
RT   "Structural relationship of human interferon alpha genes and
RT   pseudogenes.";
RL   J. Mol. Biol. [illegible] (1986).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=81107815; PubMed=[illegible];
RA   Linnane A.W., Bellamy M.W., M[illegible], Marcuccilli C.J.,
RA   Murphy M., Nisbet I.T., Nocentini G., Wood W.I.;
RT   "Nucleotide sequence of cDNA coding for human leukocyte
RT   alpha-species of interferon having antiviral activity.";
RT   Nucleic Acids Res. 9:4079-4093(1981).
RL   Biochem. Int. [illegible] (1986).
RN   [3]
RP   SEQUENCE OF 24-58.
RX   MEDLINE 98087498; PubMed=9425112;
RA   Nyman T.A., Toppila S., Parkkinen J., Kalkkinen N.;
RT   "Identification of nine interferon-alpha subtypes [illegible]
RT   virus-induced human peripheral blood leucocytes.";
RL   Biochem. J. 329:295-302(1998).
RN   [4]
RP   POLYMORPHISM.
RX   MEDLINE 97474411; PubMed=[illegible];
RA   Hussain M., Gill D.S., Liao M.-J.;
```



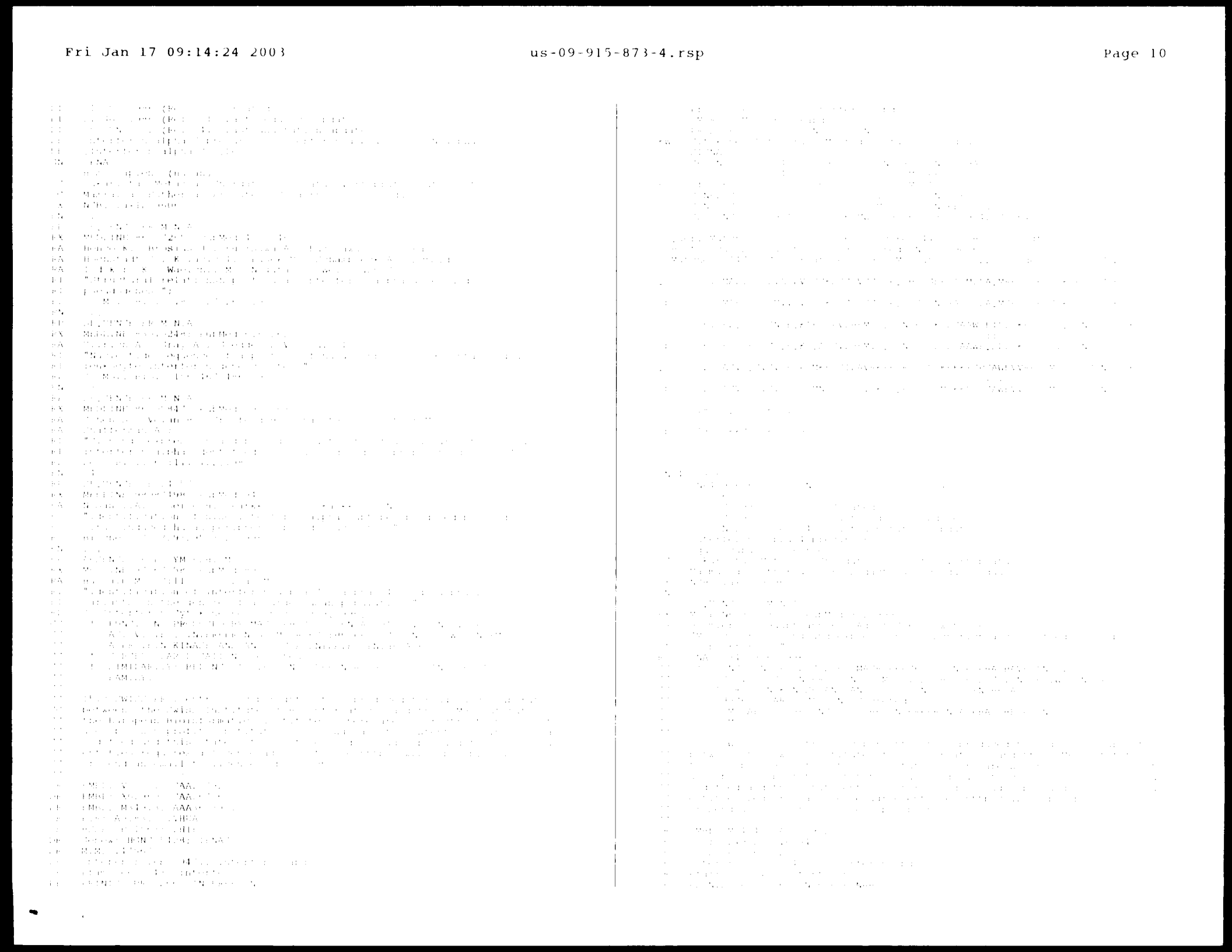


```
DR   InterPro; IPR000471; Interferon_abd.
DR   Pfam; PF00143; interferon; 1.
DR   PRINTS; PR00266; INTERFERONA.
DR   ProDom; PD000550; Interferon_abd; 1.
DR   SMART; SM00076; IFabd; 1.
DR   PROSITE; PS00252; INTERFERON_A_B_D; 1.
KW   Cytokine; Antiviral; Multigene family; Polymorphism; Signal.
FT   SIGNAL        1     23
FT   CHAIN        24    189       INTERFERON ALPHA-17.
FT   DISULFID     24    122       BY SIMILARITY.
FT   DISULFID     52    162       BY SIMILARITY.
FT   VARIANT     184    184       I -> R.
FT                                /FTId=VAR_013020.
FT   CONFLICT     57     57       H -> P (IN REF. 1).
FT   CONFLICT     78     78       S -> P (IN REF. 3).
SQ   SEQUENCE   189 AA;  21728 MW;  0448EAEAB9D7EC32 CRC64;

  Query Match             80.3%;  Score 770.5;  DB 1;  Length 189;
  Best Local Similarity   80.4%;  Pred. No. 7.6e-64;
  Matches  152;  Conservative   17;  Mismatches   15;  Indels    1;  Gaps    1;

Qy       1 MALIFALLVALLVLSCKSSCSVGCDLPQTHSLGSRRTLMLLAQMRRISLFSCLKDRHDFG 60
           |||:|:||:|:|||| || ||:|||||||||||:|| |:||||| ||| || ||||:|||
Db       1 MALSFSLLMAVLVLSYKSICSLGCDLPQTHSLGNRRALILLAQMGRISPFSCLKDRHDFG 60

Qy      61 FPQEEFDGNQFQKAPAISVLHEMIQQTFNLFSTKDSSAAWDETLLDKFYTELYQQLNDLE 120
           ||||| |||||| : | |||||||| ||||||:||||||::::|:|. ||| |:| :.|
Db      61 LPQEEFDGNQFQKTQAISVLHEMIQQIFNLFSTKDSSAAWDQSLLEKFYIELYQQLNDLE 120

Qy     120 ACVIQGVGVTETPLMKEDSILAVRKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSIN 179
           ||||| || :|||| |||||||||||||||||| |||||||||||:|:|.|| | ||
Db     121 ACVIQEVGVMETFLMNEDSILAVRKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTN 180

Qy     180 LQESLRSKE 188
           ||: || |:
Db     181 LQKILRRKD 189


RESULT 8
INAK_HUMAN
ID   INAK_HUMAN     STANDARD;      PRT:   189 AA.
AC   P01568;
DT   21-JUL-1986 (Rel. 01, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Interferon alpha-21 precursor (Interferon alpha-F) (LeIF F).
GN   IFNA21.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=81148795; PubMed=6164083;
RA   Goeddel D.V., Leung D.W., Dull T.J., Gross M., Lawn R.M.,
RA   McCandliss R., Seeburg P.H., Ullrich A., Yelverton E., Gray P.W.;
RT   "The structure of eight distinct cloned human leukocyte interferon
RT   cDNAs.";
RL   Nature 290:20-26(1981).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Gren E.Y., Berzin V.M., Isimanis A.Y., Apsalon G.E., Vishnevskii Y.I.,
RA   Yansone I.V., Dishler A.V., Pudova N.V., Smorodintsev A.A.,
RA   Iovlev V.I., Stepanov A.N., Feldmane G.Y., Meldrais Y.A., Lozha V.F.,
RA   Kavsan V.M., Efimov V.A., Sverdlov E.D.;
RT   "A new type of leukocytic interferon.";
RL   Dokl. Biochem. 269:91-95(1983).
RN   [3]
RP   SEQUENCE OF 24-58.
RX   MEDLINE=98087498; PubMed=9425112;
RA   Nyman T.A., Toeloe H., Parkkinen J., Kalkkinen N.;
RT   "Identification of nine interferon-alpha subtypes produced by Sendai
RT   virus-induced human peripheral blood leucocytes.";
```

```
DR   InterPro; IPR000471; Interferon_abd.
DR   Pfam; PF00143; interferon; 1.
DR   PRINTS; PR00266; INTERFERONA.
DR   ProDom; PD000550; Interferon_abd; 1.
DR   SMART; SM00076; IFabd; 1.
DR   PROSITE; PS00252; INTERFERON_A_B_D; 1.
KW   Cytokine; Antiviral; Multigene family; Polymorphism; Signal.
FT   SIGNAL        1     23
FT   CHAIN        24    189       INTERFERON ALPHA-17.
FT   DISULFID     24    122       BY SIMILARITY.
FT   DISULFID     52    162       BY SIMILARITY.
FT   VARIANT     184    184       I -> R.
FT                                /FTId=VAR_013020.
FT   CONFLICT     57     57       H -> P (IN REF. 1).
FT   CONFLICT     78     78       S -> P (IN REF. 3).
SQ   SEQUENCE   189 AA;  21728 MW;  0448EAEAB9D7EC32 CRC64;

  Query Match             80.3%;  Score 770.5;  DB 1;  Length 189;
  Best Local Similarity   80.4%;  Pred. No. 7.6e-64;
  Matches  152;  Conservative   17;  Mismatches   15;  Indels    1;  Gaps    1;

Qy       1 MALIFALLVALLVLSCKSSCSVGCDLPQTHSLGSRRTLMLLAQMRRISLFSCLKDRHDFG 60
           |||:|:||:|:|||| || ||:|||||||||||:|| |:||||| ||| || ||||:|||
Db       1 MALSFSLLMAVLVLSYKSICSLGCDLPQTHSLGNRRALILLAQMGRISPFSCLKDRHDFG 60

Qy      61 FPQEEFDGNQFQKAPAISVLHEMIQQTFNLFSTKDSSAAWDETLLDKFYTELYQQLNDLE 120
           ||||| |||||| : | |||||||| ||||||:||||||::::|:|. ||| |:| :.|
Db      61 LPQEEFDGNQFQKTQAISVLHEMIQQIFNLFSTKDSSAAWDQSLLEKFYIELYQQLNDLE 120

Qy     120 ACVIQGVGVTETPLMKEDSILAVRKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSIN 179
           ||||| || :|||| |||||||||||||||||| |||||||||||:|:|.|| | ||
Db     121 ACVIQEVGVMETFLMNEDSILAVRKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTN 180

Qy     180 LQESLRSKE 188
           ||: || |:
Db     181 LQKILRRKD 189


RESULT 8
INAK_HUMAN
ID   INAK_HUMAN     STANDARD;      PRT:   189 AA.
AC   P01568;
DT   21-JUL-1986 (Rel. 01, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Interferon alpha-21 precursor (Interferon alpha-F) (LeIF F).
GN   IFNA21.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=81148795; PubMed=6164083;
RA   Goeddel D.V., Leung D.W., Dull T.J., Gross M., Lawn R.M.,
RA   McCandliss R., Seeburg P.H., Ullrich A., Yelverton E., Gray P.W.;
RT   "The structure of eight distinct cloned human leukocyte interferon
RT   cDNAs.";
RL   Nature 290:20-26(1981).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Gren E.Y., Berzin V.M., Isimanis A.Y., Apsalon G.E., Vishnevskii Y.I.,
RA   Yansone I.V., Dishler A.V., Pudova N.V., Smorodintsev A.A.,
RA   Iovlev V.I., Stepanov A.N., Feldmane G.Y., Meldrais Y.A., Lozha V.F.,
RA   Kavsan V.M., Efimov V.A., Sverdlov E.D.;
RT   "A new type of leukocytic interferon.";
RL   Dokl. Biochem. 269:91-95(1983).
RN   [3]
RP   SEQUENCE OF 24-58.
RX   MEDLINE=98087498; PubMed=9425112;
RA   Nyman T.A., Toeloe H., Parkkinen J., Kalkkinen N.;
RT   "Identification of nine interferon-alpha subtypes produced by Sendai
RT   virus-induced human peripheral blood leucocytes.";
```

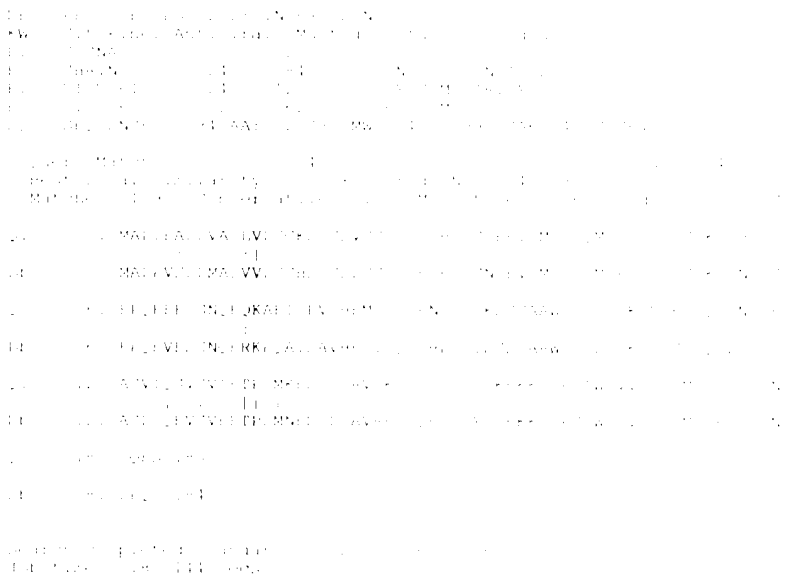


```
SQ   SEQUENCE   189 AA;  21845 MW;  CEC6B8996FDA7068 CRC64;

  Query Match             79.8%;  Score 766.5;  DB 1;  Length 189;
  Best Local Similarity   80.4%;  Pred. No. 1.8e-63;
  Matches  152;  Conservative   16;  Mismatches   20;  Indels    1;  Gaps    1;

Qy       1 MAITFALLVALLVLSCKSSCSVGCDLPQTHSLGSRRTLMLLAQMRRISLFSCLKDRHDFR 60
           |||:|:|:|:|||| || ||:|||||||||||:|| |:|| || ||| ||||||||||
Db       1 MALSFSLLMAVLVLSYKSICSLGCDLPQTHSLGNRRALILLGQMCRISPFSCLKDRHDFK 60

Qy      61 FPQEFF-GNQFQKAETIPVLHEMIQQIFNLFSTKDSSAAWDETLLDKFYIELYQQLNDLE 119
           ||||| ||||||| | ||||||||| |||||||:|||||||::|||:|| ||||||||||
Db      61 IPQEFFDGNQFQKAQAISVLHEMIQQTFNLFSTEDSSAAWDQSLLEKFSIELYQQLNDLE 120

Qy     120 ACVIQEVGVEETPLMNEDSILAVRKYFQRITLYLTEKKYSPCAWEVVRAEIMRSFSLSTN 179
           ||||| ||| ||||| |||||||||||||||||| |:|||||||||||||:|:|:|:|:| 
Db     121 ACVIQEVGVEETPLMNEDSILAVRKYFQRITLYLTEKKYSPCAWEVVRAEIMRSFSLSTN 180

Qy     180 LQESLRSKE 188
           ||: || |:
Db     181 LQKRLRRKD 189


RESULT 11
INA8_HUMAN
ID   INA8_HUMAN     STANDARD;      PRT;   189 AA.
AC   P32881; P09236; P01565;
DT   21-JUL-1986 (Rel. 01, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Interferon alpha-8 precursor (Interferon alpha-H2) (Interferon
DE   alpha-B) (LeIF B).
GN   IFNA8.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86037205; PubMed=4057246;
RA   Henco K., Brosius J., Fujisawa A., Fujisawa J.-I., Haynes J.R.,
RA   Hochstadt J., Kovacic T., Pasek M., Schamboeck A., Schmid J.,
RA   Todokoro K., Waelchli M., Nagata S., Weissmann C.;
RT   "Structural relationship of human interferon alpha genes and
RT   pseudogenes.";
RL   J. Mol. Biol. 185:227-260(1985).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=84183614; PubMed=6425903;
RA   Bowden D.W., Mao J., Gill T., Hsiao K., Lillquist J.S., Testa D.,
RA   Vovis G.F.;
RT   "Cloning of eukaryotic genes in single-strand phage vectors: the
RT   human interferon genes.";
RL   Gene 27:87-99(1984).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=81174733; PubMed=6164048;
RA   Yelverton E., Leung D., Weck P., Gray P.W., Goeddel D.V.;
RT   "Bacterial synthesis of a novel human lymphocyte interferon.";
RL   Nucleic Acids Res. 9:731-741(1981).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=81148795; PubMed=6164083;
RA   Goeddel D.V., Leung D.W., Dull T.J., Gross M., Lawn R.M.,
RA   McCandliss R., Seeburg P.H., Ullrich A., Yelverton E., Gray P.W.;
RT   "The structure of eight distinct cloned human leukocyte interferon
RT   cDNAs.";
RL   Nature 290:20-26(1981).
RN   [5]
RP   SEQUENCE OF 24-53.
RX   MEDLINE=98087498; PubMed=9425112;
RA   Nyman T.A., Lehtse M., Parkkinen J., Kalkkinen N.,
RT   "Identification of nine interferon-alpha subtypes produced by Sendai

RL   "Gene structure of human interferon alpha-8 and its pseudogenes.";
RL   Biochem. J. 320:595-602(1996).
RN   [6]
RP   ABSENCE OF GLYCOSYLATION.
RX   MEDLINE=84140669; PubMed=6607136;
RA   Hussain M., Gill D.S., Lin M., Costa G.;
RT   "Interferon-alpha MB is the only variant of interferon alpha-1
RT   identified in a large human population.";
RL   J. Interferon Cytokine Res. 14:5-8(1994).
CC   -!- FUNCTION: PRODUCED BY MANY CELLS, IN ACTION HAVE ANTIVIRAL
CC       ACTIVITIES, INTERFERON STIMULATED GENE PRODUCTION IN NORMAL
CC       AND TRANSFORMED CELLS.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: BELONGS TO THE INTERFERON ALPHA, BETA AND OMEGA
CC       FAMILY.
CC   ---------- ---------- ---------- ---------- ----------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------- ---------- ---------- ---------- ----------
DR   EMBL; V00545; CAA23806.1; -.
DR   EMBL; Reviewed; AAA52716.1; -.
DR   EMBL; X01142; CAA25511.1; -.
DR   EMBL; V00546; CAA23807.1; -.
DR   PIR; B23475; IVHUA8.
DR   PIR; A01562; IVHUA1.
DR   HSSP; P01563; 2AUE.
DR   Genew; HGNC:5422; IFNA8.
DR   MIM; 147568; -.
DR   InterPro; IPR000471; Interferon alpha/beta.
DR   Pfam; PF00143; interferon; 1.
DR   PRINTS; PR00265; INTERFERNAP.
DR   ProDom; PD000275; Interferon; 1.
DR   SMART; SM00076; IFabd; 1.
DR   PROSITE; PS00252; INTERFERON_A_B_1; 1.
KW   Cytokine; Antiviral; Multigene family; Signal.
FT   SIGNAL       1     23
FT   CHAIN       24    189       INTERFERON ALPHA-8.
FT   DISULFID    24    138       BY SIMILARITY.
FT   DISULFID    92    150       BY SIMILARITY.
FT   CONFLICT     8      8       I -> M (IN REF. 6 AND 5).
FT   CONFLICT   127    134       SVMLVQCS -> VVCQ (IN REF. 6 AND 5).
SQ   SEQUENCE   189 AA;  21845 MW;  CEC6B8996FDA7068 CRC64;

  Query Match             79.8%;  Score 766.5;  DB 1;  Length 189;
  Best Local Similarity   80.4%;  Pred. No. 1.8e-63;
  Matches  152;  Conservative   16;  Mismatches   20;  Indels    1;  Gaps    1;

Qy       1 MAITFALLVALLVLSCKSSCSVGCDLPQTHSLGSRRTLMLLAQMRRISLFSCLKDRHDFR 60
           |||:|:|:|:|||| || ||:|||||||||||:|| |:|| || ||| ||||||||||
Db       1 MALSFSLLMAVLVLSYKSICSLGCDLPQTHSLGNRRALILLGQMCRISPFSCLKDRHDFK 60

Qy      61 FPQEFF-GNQFQKAETIPVLHEMIQQIFNLFSTKDSSAAWDETLLDKFYIELYQQLNDLE 119
           ||||| ||||||| | ||||||||| |||||||:|||||||::|||:|| ||||||||||
Db      61 IPQEFFDGNQFQKAQAISVLHEMIQQTFNLFSTEDSSAAWDQSLLEKFSIELYQQLNDLE 120

Qy     120 ACVIQEVGVEETPLMNEDSILAVRKYFQRITLYLTEKKYSPCAWEVVRAEIMRSFSLSTN 179
           ||||| ||| ||||| |||||||||||||||||| |:|||||||||||||:|:|:|:|:| 
Db     121 ACVIQEVGVEETPLMNEDSILAVRKYFQRITLYLTEKKYSPCAWEVVRAEIMRSFSLSTN 180

Qy     180 LQESLRSKE 188
           ||: || |:
Db     181 LQKRLRRKD 189


RESULT 12
INA7_HUMAN
ID   INA7_HUMAN     STANDARD;      PRT;   189 AA.
AC   P01567; Q13067;
```



```
DR   ProDom; PD000550; Interferon_abd; 1.
DR   SMART; SM00076; IFabd; 1.
DR   PROSITE; PS00252; INTERFERON_A_B_D; 1.
KW   Cytokine; Antiviral; Multigene family; Signal.
FT   SIGNAL        1     23
FT   CHAIN        24    184       INTERFERON ALPHA-4.
FT   DISULFID     24    122       BY SIMILARITY.
FT   DISULFID     52    162       BY SIMILARITY.
SQ   SEQUENCE   184 AA;  20860 MW;  FC087F46F11C68EC CRC64;

  Query Match             75.4%;  Score 722.5;  DB 1;  Length 184;
  Best Local Similarity   76.6%;  Pred. No. 1.4e-59;
  Matches  141;  Conservative   18;  Mismatches   24;  Indels    1;  Gaps    1;

Qy       1 MALTFALLVALLVLSCKSSCSVGCDLPQTHSLGSRRTLMLLAQMRKISLFSCLKDRHDFG 60
           ||| |:||:||:|||| |||||:|||| ||||| | |||| ||||| ||||||||:||
Db       1 MALPFSLLMALVVLSCHSSCSLGCDLPHTHSLGNTRVLMLLGQMRRISPFSCLKDRNDFG 60

Qy      61 FPQEEF-GNQFQKAETIPVLHEMIQQIFNLFSTKDSSAAWDETLLDKFYTELYQQLNDLE 119
           |||| | ||||:| : |  :|| |||||:|||  |||||||:|||| || |||| :|
Db      61 FPQEVFDGNQFRKPQAISAVHETIQQIFHLFSTDGSSAAWDESLLDKLYIELYQQLTDLE 120

Qy     120 ACVIQGVGVTETPLMKEDSILAVRKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTN 179
           ||: | ||| ||||| |||:|||:|||||||||:|||||||||:|||| |||| |||
Db     121 ACLNQEVGVEEIPLMNEDSILAVRKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTN 180

Qy     180 LQES 183
           |  :|
Db     181 LPQS 184


RESULT 14
INA2_HORSE
ID   INA2_HORSE     STANDARD;      PRT;   184 AA.
AC   P05004;
DT   13-AUG-1987 (Rel. 05, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Interferon alpha-2 precursor.
OS   Equus caballus (Horse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Perissodactyla; Equidae; Equus
OX   NCBI_TaxID=9796;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87053170; PubMed=3022999;
RA   Himmler A., Hauptmann R., Adolf G.R., Svetly P.;
RT   "Molecular cloning and expression in Escherichia coli of equine type
RT   1 interferons.";
RL   DNA 5:345-356(1986).
CC   -!- FUNCTION: PRODUCED BY MACROPHAGES, IFN ALPHA HAVE ANTIVIRAL
CC       ACTIVITIES. INTERFERON STIMULATES THE PRODUCTION OF TWO ENZYMES:
CC       A PROTEIN KINASE AND AN OLIGOADENYLATE SYNTHETASE.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: BELONGS TO THE INTERFERON ALPHA, BETA AND DELTA
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M14541; AAA30950.1; -.
DR   EMBL; A15987; CAA01258.1; -.
DR   PIR; B24912; IVHOA2.
DR   HSSP; P01563; 2HIE.
DR   InterPro; IPR000471; Interferon_alpha_beta
DR   Pfam; PF00143; interferon; 1.
DR   PRINTS; PR00266; INTERFERONAB.
DR   ProDom; PD000550; Interferon_abd; 1.
```

```
DR   SMART; SM00076; IFabd; 1.
DR   PROSITE; PS00252; INTERFERON_A_B_D; 1.
KW   Cytokine; Antiviral; Multigene family; Signal.
FT   SIGNAL        1     23
FT   CHAIN        24    184       INTERFERON ALPHA-4.
FT   DISULFID     24    122       BY SIMILARITY.
FT   DISULFID     52    162       BY SIMILARITY.
SQ   SEQUENCE   184 AA;  20860 MW;  FC087F46F11C68EC CRC64;

  Query Match             74.8%;  Score 716.5;  DB 1;  Length 184;
  Best Local Similarity   78.1%;  Pred. No. 3.1e-59;
  Matches  144;  Conservative   18;  Mismatches   24;  Indels    1;  Gaps    1;

Qy       1 MALTFALLVALLVLSCKSSCSVGCDLPQTHSLGSRRTLMLLAQMRKISLFSCLKDRHDFG 60
           | | |||:| |||| |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db       1 MALPFSLLMALVVLSCHSSCSLGCDLPHTHSLGNTRVLMLLGQMRRISPFSCLKDRNDFG 60

Qy      61 FPQEEF-GNQFQKAETIPVLHEMIQQIFNLFSTKDSSAAWDETLLDKFYTELYQQLNDLE 119
           |||| | ||||:| : |  :|| |||||:|||  |||||||:|||| || |||| :|
Db      61 FPQEVFDGNQFRKPQAISAVHETIQQIFHLFSTDGSSAAWDESLLDKLYIELYQQLTDLE 120

Qy     120 ACVIQGVGVTETPLMKEDSILAVRKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTN 179
           ||: | ||| ||||| |||:|||:|||||||||:|||||||||:|||| |||| |||
Db     121 ACLNQEVGVEEIPLMNEDSILAVRKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTN 180

Qy     180 LQES 183
           |  :|
Db     181 LPQS 184


RESULT 14
INA2_HORSE
ID   INA2_HORSE     STANDARD;      PRT;   184 AA.
AC   P05004;
DT   13-AUG-1987 (Rel. 05, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Interferon alpha-2 precursor.
OS   Equus caballus (Horse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Perissodactyla; Equidae; Equus
OX   NCBI_TaxID=9796;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87053170; PubMed=3022999;
RA   Himmler A., Hauptmann R., Adolf G.R., Svetly P.;
RT   "Molecular cloning and expression in Escherichia coli of equine type
RT   1 interferons.";
RL   DNA 5:345-356(1986).
CC   -!- FUNCTION: PRODUCED BY MACROPHAGES, IFN ALPHA HAVE ANTIVIRAL
CC       ACTIVITIES. INTERFERON STIMULATES THE PRODUCTION OF TWO ENZYMES:
CC       A PROTEIN KINASE AND AN OLIGOADENYLATE SYNTHETASE.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: BELONGS TO THE INTERFERON ALPHA, BETA AND DELTA
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M14541; AAA30950.1; -.
DR   EMBL; A15987; CAA01258.1; -.
DR   PIR; B24912; IVHOA2.
DR   HSSP; P01563; 2HIE.
DR   InterPro; IPR000471; Interferon_alpha_beta
DR   Pfam; PF00143; interferon; 1.
DR   PRINTS; PR00266; INTERFERONAB.
DR   ProDom; PD000550; Interferon_abd; 1.
DR   SMART; SM00076; IFabd; 1.
```



OM protein - protein search, using sw model

Run on: January 17, 2003, 06:11:19 ; Search time 55.[illegible] Seconds
(without alignments)
699.371 Million cell updates/sec

Title: US-09-915-873-4
Perfect score: 960
Sequence: 1 MALIFALLVALVVLS[illegible]KSS[illegible]..........[illegible] 189

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

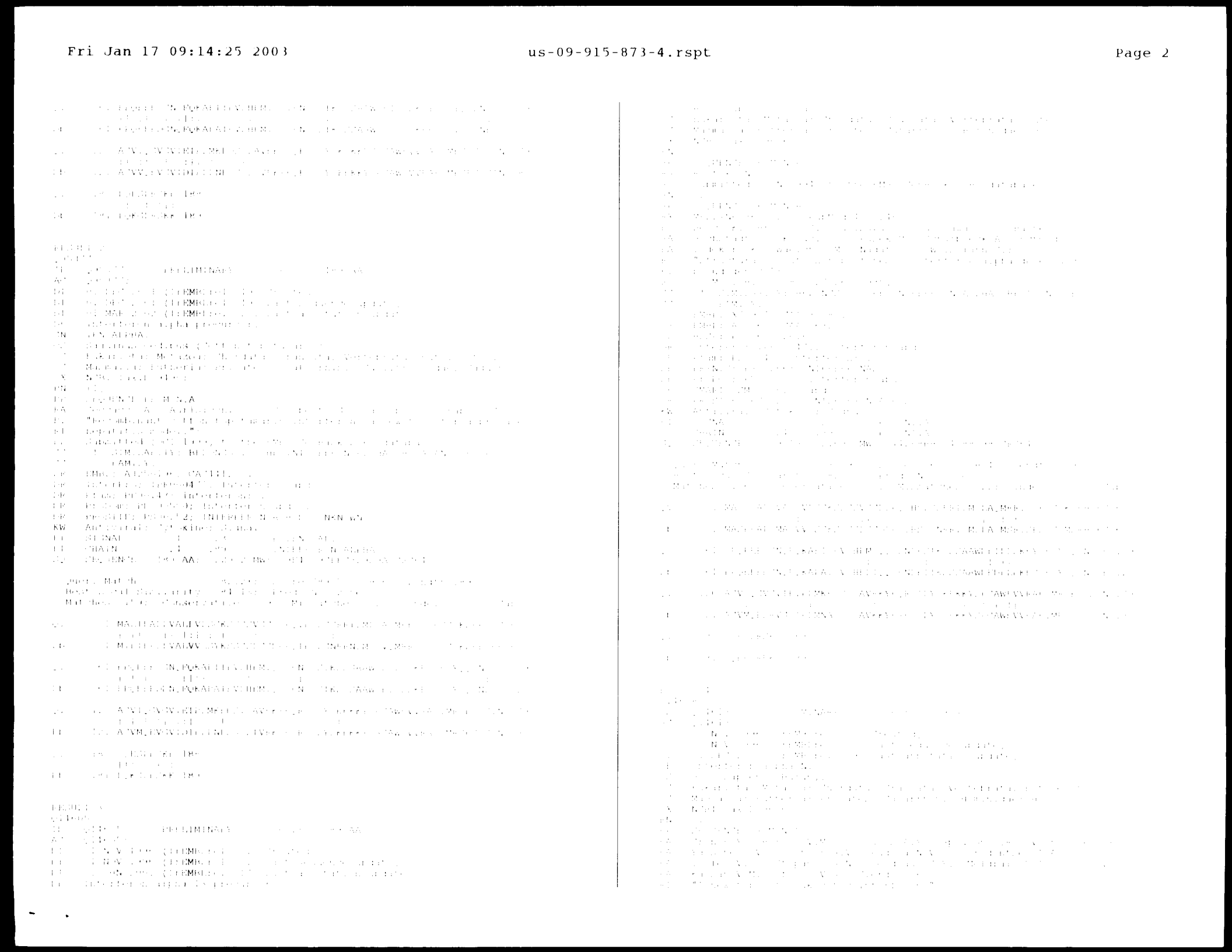
Result No.	Score	% Query Match	Length	DB	ID	Description
1	814.5	84.8	189	6	Q95J78	q95j78 saguinus oe
2	789.5	82.2	189	6	Q95J77	q95j77 saguinus oe
3	776.5	80.9	189	4	Q14605	q14605 homo sapien
4	756.5	78.8	189	4	Q14618	q14618 homo sapien
5	739.5	77.0	181	4	Q14608	q14608 homo sapien
6	688.5	71.7	166	4	Q9UMJ3	q9umj3 homo sapien
7	677.5	70.6	166	4	Q8WZ68	q8wz68 homo sapien
8	596	62.1	190	11	Q99N22	q99n22 marmota mon
9	590	61.5	190	11	Q8R5J5	q8r5j5 marmota mon
10	583	60.7	201	11	Q8R5K9	q8r5k9 marmota mon
11	580.5	60.5	190	11	Q61718	q61718 mus musculu
12	573	59.7	190	11	Q8R5J4	q8r5j4 marmota mon
13	568	59.2	190	11	Q8VHY8	q8vhy8 marmota mon
14	568	59.2	197	11	Q8R5J7	q8r5j7 marmota mon
15	567	59.1	190	11	Q8R5J3	q8r5j3 marmota mon
16	565	58.9	186	11	Q8R5J6	q8r5j6 marmota mon
17	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
18	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
19	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
20	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
21	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
22	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
23	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
24	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
25	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
26	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
27	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
28	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
29	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
30	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
31	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
32	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
33	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
34	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
35	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
36	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
37	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
38	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
39	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
40	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
41	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
42	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
43	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
44	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
45	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

ALIGNMENTS

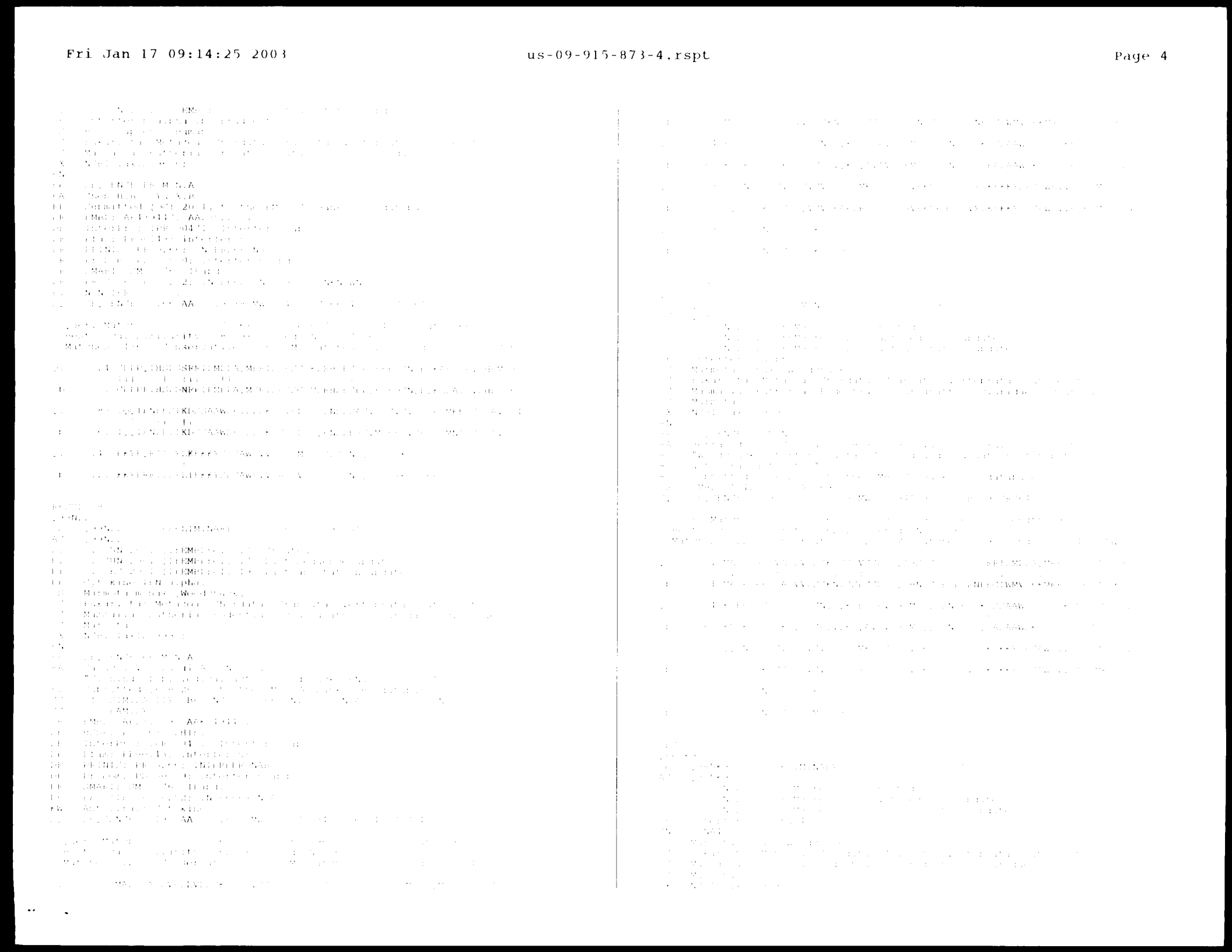
RESULT 1
Q95J78
ID Q95J78 PRELIMINARY; PRT; 189 AA.
AC Q95J78;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Interferon alpha precursor.
GN IFN ALPHA.
OS Saguinus oedipus (Cotton-top tamarin).
OC Eukaryota; Metazoa; Chordata; [illegible]
OC Mammalia; Eutheria; Primates; [illegible]
OX NCBI_TaxID=94[illegible];
RN [1]
RP SEQUENCE FROM N.A.
RA [illegible]
RT "Recombinant cotton-top tamarin interferon [illegible]
RT hepatitis model.";
RL Submitted ([illegible]) to the EMBL/GenBank/DDBJ databases.
CC [illegible] SIMILARITY: BELONGS TO THE INTERFERON ALPHA, BETA AND [illegible]
CC FAMILY.
DR EMBL; [illegible]
DR InterPro; IPR[illegible]; Interferon [illegible]
DR Pfam; PF00143; Interferon; 1.
DR PRINTS; PR[illegible]; Interferon [illegible]
DR PROSITE; PS[illegible]; INTERFERON_A_B_D; [illegible]
KW Antiviral; Cytokine; Signal.
FT SIGNAL 1 [illegible] POTENTIAL.
FT CHAIN [illegible] 189 [illegible] POTENTIAL.
SQ SEQUENCE 189 AA; [illegible] MW; [illegible] CRC64;

Query Match [illegible]; Score [illegible]; DB [illegible]; Length 189;
Best Local Similarity [illegible]; Pred. No. [illegible];
Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

Qy 1 MALIFALLVALVVLS[illegible]KS[illegible]SV[illegible]LHSL[illegible]SRR[illegible]LMLLA[illegible]MRRIS[illegible]
Db 1 MALIF[illegible]LVALVVLSYKSS[illegible]LHSL[illegible]NRR[illegible]LMLLA[illegible]MRRIS[illegible]

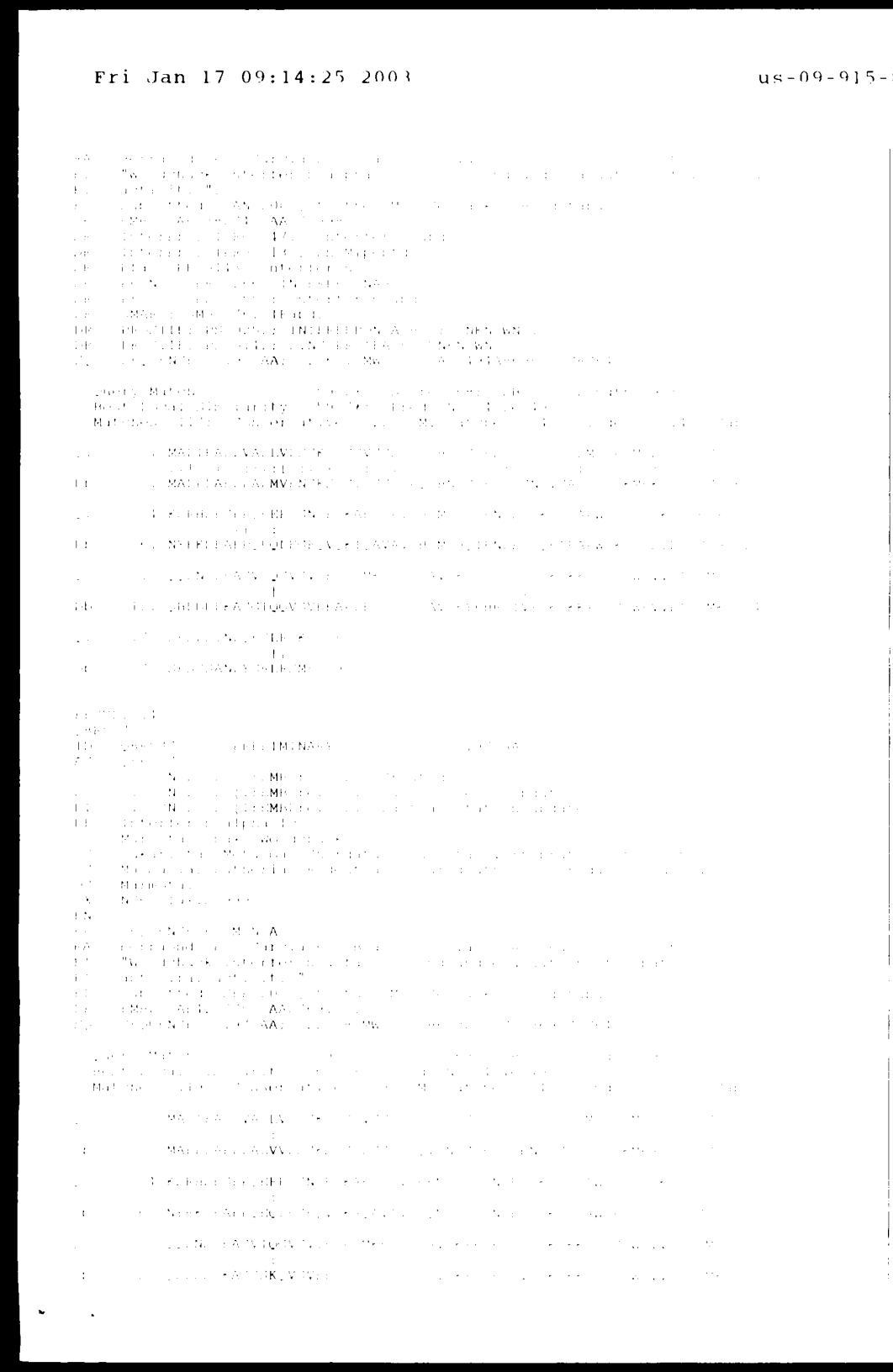


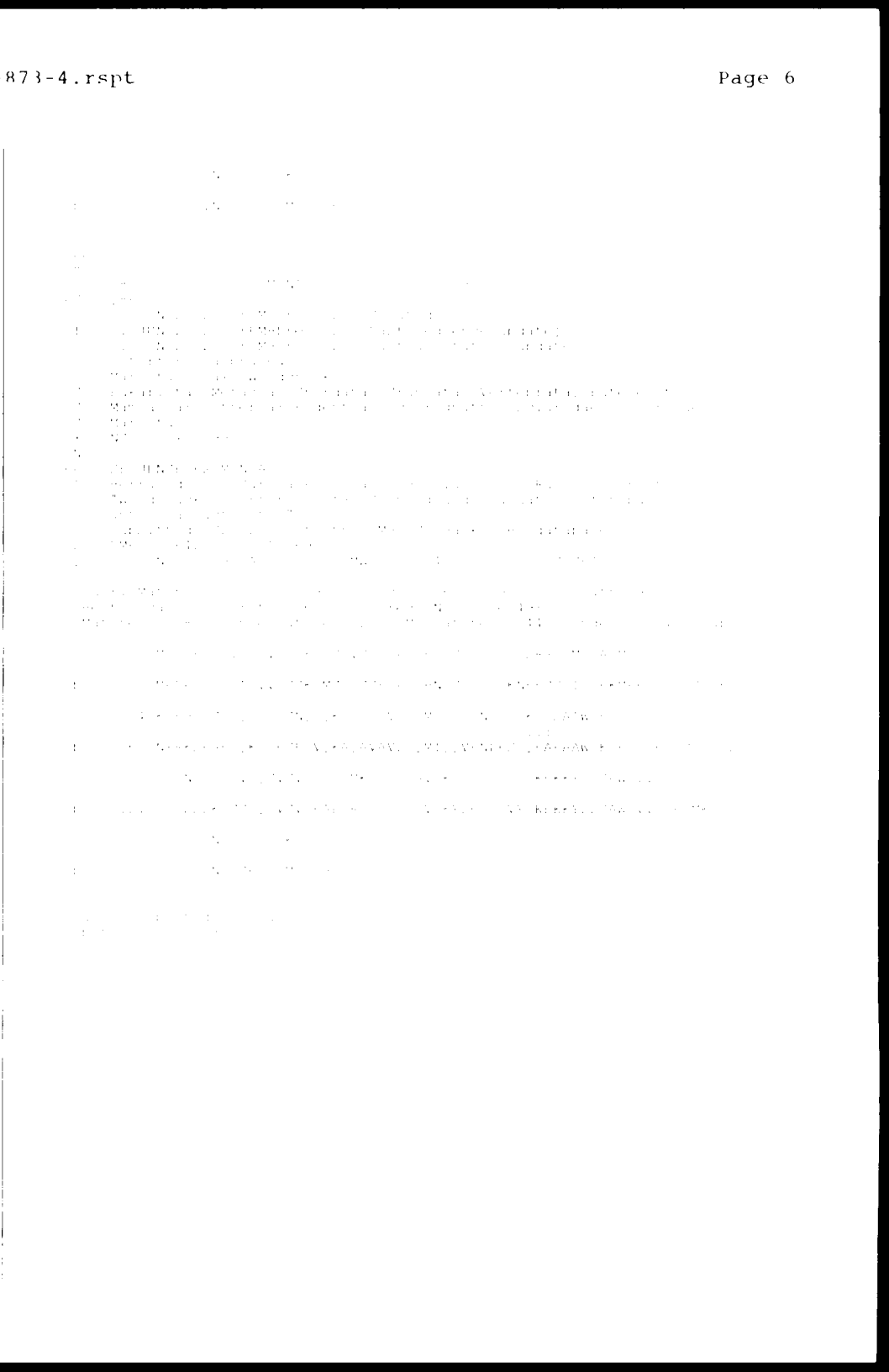
```
RL   Dokl. Biochem. 269:91-95(1983).
CC   -!- SIMILARITY: BELONGS TO THE INTERFERON ALPHA, BETA AND DELTA
CC       FAMILY.
DR   EMBL; X00140; CAA24970.1; -.
DR   HSSP; P01563; 2HIE.
DR   InterPro; IPR000471; Interferon_abd.
DR   Pfam; PF00143; interferon; 1.
DR   PRINTS; PR00266; INTERFERONAB.
DR   ProDom; PD000550; Interferon_abd; 1.
DR   SMART; SM00076; IFabd; 1.
DR   PROSITE; PS00252; INTERFERON_A_B_D; 1.
KW   Antiviral; Cytokine.
SQ   SEQUENCE   189 AA;  21810 MW;  B6D8D9E726E04344 CRC64;

  Query Match             78.8%;  Score 756.5;  DB 4;  Length 189;
  Best Local Similarity   79.9%;  Pred. No. 5.2e-68;
  Matches  151;  Conservative   12;  Mismatches   25;  Indels    1;  Gaps    1;

Qy       1 MALTFALLVALLVLSCKSSCSVGCDLPQTHSLGSRRTLMLLAQMRRISLFSCLKDRHDFG 60
           | ||:||:|:|||| || ||:|||||||||:||  :||||| ||| ||||||:|||
Db       1 MPLSFSLLMAVLVLSYKSICSLGCDLLPQTHSLGNKRAWTLLAQMGRISLFSCLKDRKDFR 60

Qy      61 FPQEFF-GNQFQKAETIPVLHEMIQQIFNLFSTKDSSAAWDETLLERFYIELFQQLNDLE 119
           |||| | |||||||: |  |||||| ||||||||||||||||||||||' ||:||.|||
Db      61 FPQEVFDGNQFQKAQAISAFHEMIQQTFNLFSTKDSSAAWDETLLERFYIELFQQLNDLE 120

Qy     120 ACVIQGVGVTETPLMKEDSILAVRKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTN 179
           ||| | ||| |  || ||||||||||||||||||  ||||||||||||||||||||||||| |||
Db     121 ACVTQEVGVMETALMNEDSILAVKKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTN 180

Qy     180 LQESLRSKE 188
           ||: || |:
Db     181 LQKGLRRKD 189


RESULT 5
Q14608
ID   Q14608      PRELIMINARY;      PRT;   181 AA.
AC   Q14608;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Leukocyte interferon-alpha.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE 85056523; PubMed 6548765;
RA   Gren E., Borzin V.M., Jansone I., Isimanis A., Vishnevsky Y.,
RA   Apsalons U.;
RT   "Novel human leukocyte interferon subtype and structural comparison of
RT   alpha interferon genes.";
RL   J. Interferon Res. 4:609-617(1984).
CC   -!- SIMILARITY: BELONGS TO THE INTERFERON ALPHA, BETA AND DELTA
CC       FAMILY.
DR   EMBL; M28586; AAA36041.1; -.
DR   HSSP; P01563; 2HIE.
DR   InterPro; IPR000471; Interferon_abd.
DR   Pfam; PF00143; interferon; 1.
DR   PRINTS; PR00266; INTERFERONAB.
DR   ProDom; PD000550; Interferon_abd; 1.
DR   SMART; SM00076; IFabd; 1.
DR   PROSITE; PS00252; INTERFERON_A_B_D; 1.
KW   Antiviral; Cytokine.
SQ   SEQUENCE   181 AA;  20878 MW;  4DB45120764EBAB0 CRC64;

  Query Match             77.0%;  Score 739.5;  DB 4;  Length 181;
  Best Local Similarity   81.8%;  Pred. No. 2.5e-66;
  Matches  148;  Conservative   11;  Mismatches   21;  Indels    1;  Gaps    1;

Qy       6 VALTVLSCKSSCSVGCDLPQTHSLGSRRTLMLLAQMRRISLFSCLKDRHDFGFPQEFF- 64
           |||:|||  |||||||||||||||||||||||||||||||||||||||||||||||| |
Db       1 MAVLVLSYKSICSLGCDLLPQTHSLGNKRAWTLLAQMGRISLFSCLKDRKDFRFPQEVFD 60

Qy      65 GNQFQKAETIPVLHEMIQQIFNLFSTKDSSAAWDETLLERFYIELFQQLNDLEACVIQGV 124
           |||||||| :|||||||||| ||  |  ||||||||||||||||||||||||||||||||
Db      61 GNQFQKAQAISAFHEMIQQTFNLFSTKDSSAAWDETLLERFYIELFQQLNDLEACVIQEV 120

Qy     125 GVTETPLMKEDSILAVRKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTNLQESLRS 184
           |||| ||||||||:|||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GVMETALMNEDSILAVKKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTNLQKGLRR 180

Qy     185 KE 188
           ||
Db     181 KD 182


RESULT 6
Q9UMJ4
ID   Q9UMJ4      PRELIMINARY;      PRT;   188 AA.
AC   Q9UMJ4;
DT   01-MAY-2000 (TrEMBLrel. 15, Created)
DT   01-MAY-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE   IFNA protein (Fragment).
GN   IFNA.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE 99120341; PubMed 9914503;
RA   Weber B., Weissenbach J.;
RT   "Formation of cDNA libraries specific for differentiated cells by
RT   between-related, cloned genes in E. coli.";
RL   Nucleic Acids Res. 11:5661-5676(1983).
CC   -!- SIMILARITY: BELONGS TO THE INTERFERON ALPHA, BETA AND DELTA
CC       FAMILY.
DR   EMBL; M29894; AAA75714.1; -.
DR   HSSP; P01563; 2HIE.
DR   InterPro; IPR000471; Interferon_abd.
DR   Pfam; PF00143; interferon; 1.
DR   PRINTS; PR00266; INTERFERONAB.
DR   ProDom; PD000550; Interferon_abd; 1.
DR   SMART; SM00076; IFabd; 1.
DR   PROSITE; PS00252; INTERFERON_A_B_D; 1.
KW   Antiviral; Cytokine.
FT   NON_TER       1       1
SQ   SEQUENCE   188 AA;  [illegible] MW;  [illegible] CRC64;

  Query Match             [illegible]
  Best Local Similarity   [illegible]
  Matches  [illegible]

Qy      24 CDLPQTHSLGNRRTLMLLAQMRRISLFSCLKDRHDFGFPQEFFGNQFQKAETIPVLHEM 83
           ||||||||| :|||||||||| ||  |  ||||||||||||||||||||||||||||||
Db       1 CDLPETHSLDNERTLMLLAQMSRISPSSCLMDRHDFDFPQEIFDGNQFQKAQAISAFHEM 60

Qy      84 IQQIFNLFSTKDSSAAWDETLLERFYIELFQQLNDLEACVIQGVGVTETPLMKEDSILAV 143
           | |:|.|:|||| | |:|:|  |  |  |:|:|:|:|||||||||||||||||||||||
Db      61 IQQTFNLFSTKDSSAAWDETLLERFYIELFQQLNDLEACVIQEVGVMETALMNEDSILAV 120

Qy     144 KKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTNLQESLRSKE 188
           :|||:|||||| ||||||||||||||| | |:|||||| |: ||
Db     121 KKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSKNLQKGLRRKD 165


RESULT 7
Q9WZ68
ID   Q9WZ68      PRELIMINARY;      PRT;   [illegible] AA.
AC   Q9WZ68;
DT   01-MAR-2002 (TrEMBLrel. [illegible] Created)
DT   01-MAR-2002 (TrEMBLrel. [illegible] Last sequence update)
```



```
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Berraondo P., Garcia R., Blanco B., Valos A., Ruiz L., Prieto J.;
RT   "Woodchuck interferon alpha system: cloning and evaluation of antiviral
RT   activity.";
RL   Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF338273; AAL71997.2; -.
SQ   SEQUENCE   201 AA;  22594 MW;  91E3917A5F2E 7F2 "B"1.

  Query Match             60.7%;  Score 583;  DB 11;  Length 201;
  Best Local Similarity   62.2%;  Pred. No. 1.4e-50;
  Matches  122;  Conservative   19;  Mismatches   41;  Indels   11;  Gaps    1;

Qy       1 MALTFALLVALLVLSCKSSCSVGCDLPQTHSLG---    SRRTIMLLAQMRRISLFSCL 54
           |||    |:||:||||||:||:||||||  |:||          |  || :||||  |||
Db       1 MALPLPFLLALVVLSCKSTCSLGCDLPQTHNLGLRRTRTLMLLEEMRRAPPLSPCL 60

Qy      55 KDRHDFGFPQEEFGNQFQKAETIPVLHEMIQQIFNLFSTKDSSAAWDETLLDKFYTELY 114
           | || |||::  | | |||: : ||||| ||||||||||:: ||||:|:|. | | |:
Db      61 NYRFDFAFPQKEFDGSQVQKAQAVAVLHEMTQQIFNLFSTQDASAAWDKTLLDKFYIELY 120

Qy     115 QQLNDLEACVIQGVGVTETPLMKEDSILAVRKYFQRITLYLKEKKYSPCAWEVVRAEIMR 174
           |||:||:||  | ||| | ||         |||||| |||:||||||. |||| ||| |:
Db     121 QQIMDLEACVIQQVGVEDTALMK-      AVLAVRKYFHRITVYLKEKKHSPCAWEVVRAEIMR 179

Qy     175 SFSLSTNLQESLRSKE 190
           ||| | ||   ||| |
Db     180 SFSSSPNLYGRLKSME 195


RESULT 11
Q61718
ID   Q61718      PRELIMINARY;      PRT:   190 AA.
AC   Q61718;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Alpha-interferon precursor.
GN   MUIFN-ALPHA-A.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85056526; PubMed=6094686;
RA   Daugherty B., Martin-Zanca D., Kelder B., Collier K., Sommaus L.C.,
RA   Hotta K., Pestka S.;
RT   "Isolation and bacterial expression of a murine alpha leukocyte
RT   interferon gene.";
RL   J. Interferon Res. 4:635-643(1984).
CC   -!- SIMILARITY: BELONGS TO THE INTERFERON ALPHA, BETA AND DELTA
CC       FAMILY.
DR   EMBL; M28587; AAA37889.1; -.
DR   HSSP; P01563; 2HIE.
DR   InterPro; IPR000471; Interferon_abd.
DR   Pfam; PF00143; interferon; 1.
DR   PRINTS; PR00266; INTERFERONAB.
DR   ProDom; PD000550; Interferon_abd; 1.
DR   SMART; SM00076; IFabd; 1.
DR   PROSITE; PS00252; INTERFERON_A_B_D; 1.
KW   Antiviral; Cytokine; Signal.
FT   SIGNAL        1     23       POTENTIAL.
FT   CHAIN        24    190       ALPHA-INTERFERON.
SQ   SEQUENCE   190 AA;  21903 MW;  A60AD98E1DF7D33F CRC64;

  Query Match             60.5%;  Score 580.5;  DB 11;  Length 190;
  Best Local Similarity   60.4%;  Pred. No. 2.3e-50;
  Matches  114;  Conservative   29;  Mismatches   45;  Indels    1;  Gaps    1;

Qy       1 MALTFALLVALLVLSCKSSCSVGCDLPQTHSLGSRRTLMLLAQMRRISLFSCLKDRHDFG 60
           ||    | |: |||:|  |:||:|||||||:| :|| | || ||||:|  |.| || |
Db       1 MARLCAFLMTLLVMSYWSTCSLGCDLPQTHNLRNKRALILLVQMRRLSPLSCLKDRKDFR 60
```

```
Qy      61 FPQEEFDGNQFQKAETIPVLHEMIQQIFNLFSTKDSSAAWDETLLDKFYTELYQQLNDLE 120
           |:|:||  | |||  | | |||: : ||||| ||||||||||:: ||||:|:|. | | |:
Db      61 FPKVEVEQQQVQKAEAIPVMHELIQQTFNLFSTEHSSAAWNMTLLEQLCTGLHQQINHLQ 120

Qy     121 ACVIQGVGVTETPLMKEDSILAVRKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTN 180
           ||:||:||  | ||| | ||         |||||| |||:||||||. |||| ||| |:| |
Db     121 ACMVQEVGVMETPLMNVDSILAVKKYFRRITLYLTEKKYSPCAWEVVRVEIMRSFSFSTN 180

Qy     181 LQESLRSKE 190
           | |:|:|:|
Db     181 LNERLRRKE 189


RESULT 12
Q8MCJ4
ID   Q8MCJ4      PRELIMINARY;      PRT:   190 AA.
AC   Q8MCJ4;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Interferon alpha 5.
GN   Marmota monax (Woodchuck).
OS   Marmota monax (Woodchuck).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Sciuridae; Marmotini;
OC   Marmota.
OX   NCBI_TaxID=9995;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Berraondo P., Garcia R., Blanco B., Valos A., Ruiz L., Prieto J.;
RT   "Woodchuck interferon alpha system: cloning and evaluation of antiviral
RT   activity.";
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF338273; AAL71997.2; -.
SQ   SEQUENCE   190 AA;  21903 MW;  A60AD98E1DF7D33F CRC64;

  Query Match             60.5%;  Score 580.5;  DB 11;  Length 190;
  Best Local Similarity   60.4%;  Pred. No. 2.3e-50;
  Matches  114;  Conservative   29;  Mismatches   45;  Indels    1;  Gaps    1;

Qy       1 MALTFALLVALLVLSCKSSCSVGCDLPQTHSLGSRRTLMLLAQMRRISLFSCLKDRHDFG 60
           ||    | |: |||:|  |:||:|||||||:| :|| | || ||||:|  |.| || |
Db       1 MARLCAFLMTLLVMSYWSTCSLGCDLPQTHNLRNKRALILLVQMRRLSPLSCLKDRKDFR 60

Qy      61 FPQEEFDGNQFQKAETIPVLHEMIQQIFNLFSTKDSSAAWDETLLDKFYTELYQQLNDLE 120
           |:|:||  | |||  | | |||: : ||||| ||||||||||:: ||||:|:|. | | |:
Db      61 FPKVEVEQQQVQKAEAIPVMHELIQQTFNLFSTEHSSAAWNMTLLEQLCTGLHQQINHLQ 120

Qy     121 ACVIQGVGVTETPLMKEDSILAVRKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTN 180
           ||:||:||  | ||| | ||         |||||| |||:||||||. |||| ||| |:| |
Db     121 ACMVQEVGVMETPLMNVDSILAVKKYFRRITLYLTEKKYSPCAWEVVRVEIMRSFSFSTN 180

Qy     181 LQESLRSKE 190
           | |:|:|:|
Db     181 LNERLRRKE 189


RESULT 13
Q9VHY8
ID   Q9VHY8      PRELIMINARY;      PRT:   190 AA.
AC   Q9VHY8;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Interferon alpha 5.
GN   IFNA5.
OS   Marmota monax (Woodchuck).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Sciuridae; Marmotini;
OC   Marmota.
OX   NCBI_TaxID=9995;
RN   [1]
RP   SEQUENCE FROM N.A.
```





GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: January 17, 2003, 06:54:07 ; Search time 15.473 Seconds
(without alignments)
649.210 Million cell updates/sec

Title: US-09-915-873-4
Perfect score: 960
Sequence: 1 MALTTALLVALLVLSCKSSC.....EIMRSFSLSINLAMKLRSKE 188

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep *
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep *
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep *
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep *
5: /cgn2_6/ptodata/1/iaa/PCT_COMB.pep *
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	957	99.7	188	4	US-09-206-903A-7	Sequence 7, Appli
2	957	99.7	188	4	US-09-202-122-7	Sequence 7, Appli
3	957	99.7	188	4	US-09-206-935-9	Sequence 9, Appli
4	943	98.2	189	4	US-09-206-936-9	Sequence 9, Appli
5	937.5	97.7	189	1	US-08-026-758-4	Sequence 4, Appli
6	937	97.6	219	6	5310729-4	Patent No. 5310729
7	934.5	97.3	189	1	US-08-026-758-5	Sequence 5, Appli
8	929.5	96.8	189	6	5510472-7	Patent No. 5510472
9	895	93.2	195	6	5198345-17	Patent No. 5198345
10	851.5	88.7	188	1	US-08-249-671A-11	Sequence 11, Appl
11	851	88.6	165	1	US-08-024-330-1	Sequence 1, Appli
12	851	88.6	165	1	US-07-952-840-1	Sequence 1, Appli
13	851	88.6	165	1	US-08-456-021-1	Sequence 1, Appli
14	851	88.6	165	5	PCT-US94-01729-1	Sequence 1, Appli
15	851	88.6	198	4	US-09-333-348B-11	Sequence 11, Appl
16	845	88.0	165	4	US-09-216-500-1	Sequence 1, Appli
17	845	88.0	165	4	US-09-730-464-1	Sequence 1, Appli
18	843	87.8	165	1	US-08-249-671A-5	Sequence 5, Appli
19	840.5	87.6	166	3	US-08-819-238A-1	Sequence 1, Appli
20	840.5	87.6	166	4	US-09-479-434-1	Sequence 1, Appli
21	835	87.0	165	6	5210829-4	Patent No. 5210829
22	833.5	86.8	166	3	US-08-819-238A-13	Sequence 13, Appl
23	833.5	86.8	166	4	US-09-479-434-13	Sequence 13, Appl
24	829.5	86.4	166	1	US-08-362-453-9	Sequence 9, Appli
25	815	84.9	165	1	US-08-362-453-12	Sequence 12, Appl
26	810	84.4	165	1	US-08-362-453-14	Sequence 14, Appl
27	809.5	84.3	199	4	US-09-206-945-12	Sequence 12, Appl
28	808.5	84.2	165	1	US-08-249-671A-12	Sequence 12, Appl
29	808.5	84.2	189	1	US-08-026-758-6	Sequence 6, Appli
30	808	84.2	165	6	5256754-4	Patent No. 5256754
31	802.5	83.6	189	1	US-08-026-758-7	Sequence 7, Appli
32	802.5	83.6	189	1	US-08-026-758-8	Sequence 8, Appli
33	802.5	83.6	189	4	US-09-206-905-11	Sequence 11, Appl
34	799.5	83.3	189	1	US-08-026-758-9	Sequence 9, Appli
35	791.5	82.4	189	1	US-08-026-758-10	Sequence 10, Appl
36	791.5	82.4	189	4	US-09-206-905-10	Sequence 10, Appl
37	791.5	82.4	189	4	US-09-206-905-12	Sequence 12, Appl
38	774.5	80.7	189	1	US-08-026-758-11	Sequence 11, Appl
39	774.5	80.7	189	4	US-09-206-905-11	Sequence 11, Appl
40	774.5	80.7	189	4	US-09-206-905-13	Sequence 13, Appl
41	771.5	80.3	189	4	US-09-206-905-10	Sequence 10, Appl
42	770.5	80.3	189	1	US-08-026-758-12	Sequence 12, Appl
43	770.5	80.3	189	1	US-08-026-758-13	Sequence 13, Appl
44	770.5	80.3	189	1	US-08-026-758-14	Sequence 14, Appl
45	769.5	80.2	189	1	US-08-026-758-15	Sequence 15, Appl

ALIGNMENTS

RESULT 1
US-09-206-903A-7
; Sequence 7, Application US/09206903A
; Patent No. 6200760
; GENERAL INFORMATION:
; APPLICANT: Chen, Jian
; APPLICANT: Grabowski, Paul S.
; APPLICANT: Wu, Wilson
; TITLE OF INVENTION: Novel Human Kinase
; FILE REFERENCE: PI0424.PF1
; CURRENT APPLICATION NUMBER: US/09/206,903A
; CURRENT FILING DATE: 1998-12-7
; PRIOR APPLICATION NUMBER: US 60/068,465
; PRIOR FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 12
; SEQ ID NO 7
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-206-903A-7

Query Match 99.7%; Score 957; DB 4; Length 188;
Best Local Similarity 99.5%; Pred. No. 1.5e-103;
Matches 187; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

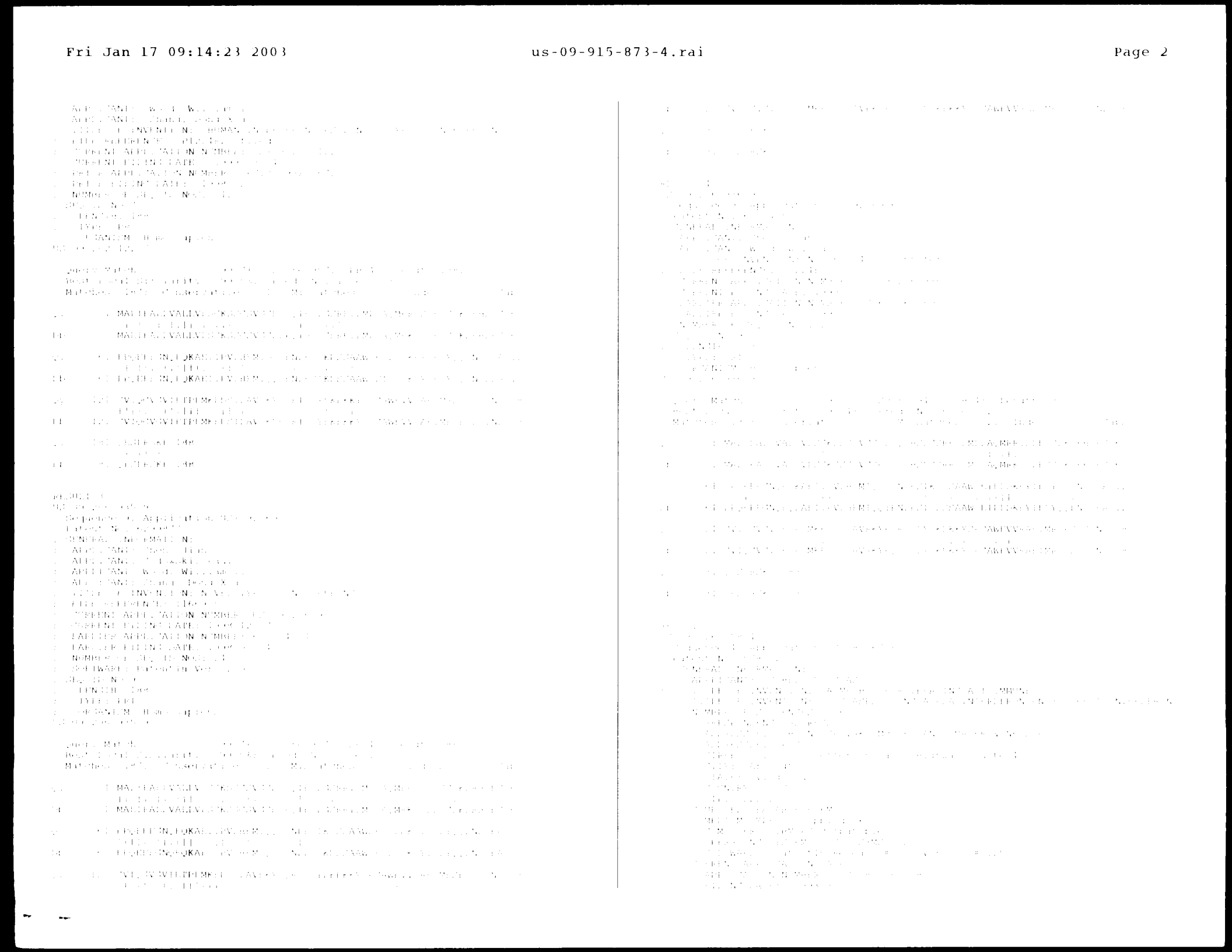
Qy 1 MALTTALLVALLVLSCKSSCSVGCDLPQTHSLGSRRTLMLLAQMRKISLFSCLKDRHDF 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MALTTALLVALLVLSCKSSCSVGCDLPQTHSLGSRRTLMLLAQMRKISLFSCLKDRHDF 60

Qy 61 DFPQEEFDNSQFQKAEAISVLHEMIQQTFNLFSTKDSSAAWDETLLDKFYTELYQQLNDL 120
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61 DFPQEEFDNSQFQKAEAISVLHEMIQQTFNLFSTKDSSAAWDETLLDKFYTELYQQLNDL 120

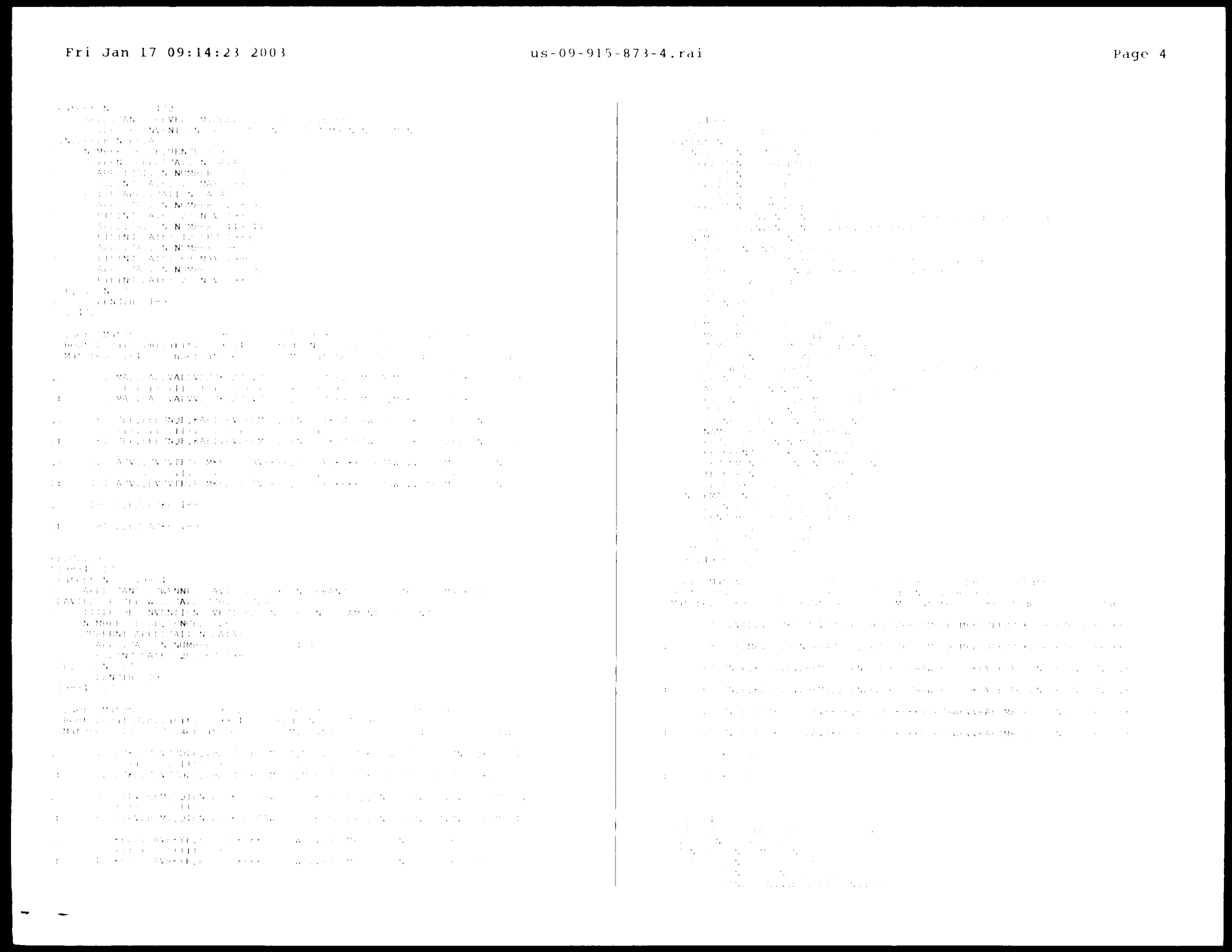
Qy 121 EACVIQEVGVEETPLMNEDSILAVKKYFQRITLYLTEKKYSPCAWEVVRAEIMRSFSLSIN 180
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 EACVIQEVGVEETPLMNEDSILAVKKYFQRITLYLTEKKYSPCAWEVVRAEIMRSFSLSIN 180

Qy 181 LAMKLRSKE 189
|||||||||
Db 181 LAMKLRSKE 189

RESULT 2
US-09-202-122-7
; Sequence 7, Application US/09202122
; Patent No. 6294669
; GENERAL INFORMATION:
; APPLICANT: Chen, Jian
; APPLICANT: Grabowski, Paul



```
;       CLASSIFICATION:  424
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Oblon, No. 5780021man F.
;       REGISTRATION NUMBER:  24,618
;       REFERENCE/DOCKET NUMBER:  1126-096-0
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (703) 413-4000
;       TELEFAX:  (703) 413-2220
;       TELEX:  248855 OPAT UR
;   INFORMATION FOR SEQ ID NO:  4.
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  189 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  unknown
;     MOLECULE TYPE:  protein
;     FEATURE:
;       NAME/KEY:  Protein
;       LOCATION:  24..189
;       OTHER INFORMATION:  /note= "IFN-alpha 2"
US-08-026-758-4

  Query Match             97.7%;  Score 937.5;  DB 1;  Length 189;
  Best Local Similarity   98.4%;  Prod. No. 2.9e-101;
  Matches  186;  Conservative    0;  Mismatches    2;  Indels    1;  Gaps    1;

Qy       1 MALTFALLVALLVLSCKSSCSVGCDLPQTHSLGSRRTLMLLAQMRKISLFSCLKDRHDFG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MALTFALLVALLVLSCKSSCSVGCDLPQTHSLGSRRTLMLLAQMRKISLFSCLKDRHDFG 60

Qy      61 FPQEEF GNQFQKAETIPVLHEMIQQIFNLFSTKDSSAAWLETLLEKFYTELYQQLNDLE 120
           |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 FPQFFPGNQFQKAETIPVLHEMIQQIFNLFSTKDSSAAWDETLLDKFYTELYQQLNDLE 120

Qy     121 ACVIQGVGVTETPLMKEDSILAVRKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTNL 180
           ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 ACVMQRVGVTETPLMKEDSILAVRKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTNL 180

Qy     181 QESLRSKE 188
           ||||||||
Db     181 QESLRSKE 189


RESULT 6
5310729-4
;Patent No. 5310729
;     APPLICANT: LERNHARDT, WALDEMAR
;     TITLE OF INVENTION: INTERFERON-RELATED POLYPEPTIDES AS "A2
;LIGANDS AND THEIR USE FOR MODULATING IMMUNE FUNCTIONS"
;     NUMBER OF SEQUENCES: 47
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/07/512,118
;       FILING DATE: 20-APR-1990
;SEQ ID NO:4:
;       LENGTH: 219
5310729-4

  Query Match             97.6%;  Score 937;  DB 6;  Length 219;
  Best Local Similarity   97.9%;  Prod. No. 4e-101;
  Matches  184;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    1;

Qy       1 MALTFALLVALLVLSCKSSCSVGCDLPQTHSLGSRRTLMLLAQMRKISLFSCLKDRHDFG 60
           ||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db      32 MALTFALLVALLGLSCKSSCSVGCDLPQTHSLGSRFTLMLLAQMRKISLFSCLKDRHDFG 91

Qy      61 FPQEEFGNQFQKAETIPVLHEMIQQIFNLFSTKDSSAAWDETLLDKFYTELYQQLNDLEA 120
           |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db      92 FPQFFPGNQFQKAETSPVLHEMIQQIFNLFSTKDSSAAWDETLLDKFYTELYQQLNDLEA 151

Qy     121 CVIQGVGVTETPLMKEDSILAVRKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTNL 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     152 CVIQGVGVTETPLMKEDSILAVRKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTNL 211

Qy     181 QESLRSKE 188
           ||||||||
Db     212 QESLRSKE 219


RESULT 7
US-08-026-758-5
; Sequence 5, Application US/ = 2
; Patent No. 5780021
; GENERAL INFORMATION:
;     APPLICANT:  SOBEL, DOUGLAS G.
;     TITLE OF INVENTION:  A METHOD FOR TREATING AUTOIMMUNE
;     TITLE OF INVENTION:  DISEASES BEING AUTOIMMUNITY AND/OR MULTIPLE M
;     NUMBER OF SEQUENCES:  20
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTA
;       ADDRESSEE:  E.C.
;       STREET:  1755 J, Jefferson Davis Highway, Suite 4
;       CITY:  Arlington
;       STATE:  Virginia
;       COUNTRY:  U.S.A.
;       ZIP:  22202
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-D
;       SOFTWARE:  Pa  Ve  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/07/
;       FILING DATE:  19930505
;       CLASSIFICATION:  424
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Oblon, Spivak, McC
;       REGISTRATION NUMBER:  24,618
;       REFERENCE/DOCKET NUMBER:  1126-
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (703) 413-4000
;       TELEFAX:  (703) 413-2220
;       TELEX:  248855 OPAT UR
;   INFORMATION FOR SEQ ID NO:  5:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  189 amino acid
;       TYPE:  amino acid
;       TOPOLOGY:  unknown
;     MOLECULE TYPE:  protein
;     FEATURE:
;       NAME/KEY:  Protein
;       LOCATION:  24..189
;       OTHER INFORMATION:  /note= "IFN-alpha A"
US-08-026-758-5

  Query Match             97.7%;  Score 937.5;  DB 1;  Length 189;
  Best Local Similarity   98.4%;  Prod. No. 2.9e-101;
  Matches  186;  Conservative    0;  Mismatches    2;  Indels    1;  Gaps    1;

Qy       1 MALTFALLVALLVLSCKSSCSVGCDLPQTHSLGSRRTLMLLAQMRKISLFSCLKDRHDFG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MALTFALLVALLVLSCKSSCSVGCDLPQTHSLGSRRTLMLLAQMRKISLFSCLKDRHDFG 60

Qy      61 FPQEEF GNQFQKAETIPVLHEMIQQIFNLFSTKDSSAAWLETLLEKFYTELYQQLNDLE 120
           |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 FPQFFPGNQFQKAETIPVLHEMIQQIFNLFSTKDSSAAWDETLLDKFYTELYQQLNDLE 120

Qy     121 ACVIQGVGVTETPLMKEDSILAVRKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTNL 180
           ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 ACVMQRVGVTETPLMKEDSILAVRKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTNL 180

Qy     181 QESLRSKE 188
           ||||||||
Db     181 QESLRSKE 189


RESULT 8
5310472-7
```



; APPLICANT: Nagabhushan, Tattanahalli
; APPLICANT: Tindall, Stephen
; APPLICANT: Bruza, Alan
; TITLE OF INVENTION: Metal-Interferon-Alfa Crystals
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schering Plough Corporation
; STREET: One Giralda Farms
; CITY: Madison
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07940
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 6.0.5
; SOFTWARE: Microsoft Word 4.00B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/[illegible]
; FILING DATE: 19930225
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA: No. 5441734e
; ATTORNEY/AGENT INFORMATION:
; NAME: Lunn, Paul G.
; REGISTRATION NUMBER: 32,743
; REFERENCE/DOCKET NUMBER: J99287
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 822 7255
; TELEFAX: 201 822 7039
; TELEX: 219165
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 165 amino acid residues
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-024-330-1

Query Match 88.6%; Score 851; DB 1; Length 165;
Best Local Similarity 100.0%; Pred. No. 2.8e-91;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 24 CDLPQTHSLGSRRTLMLLAQMRRISLFSCLKDRHDFGFPQEEFGNQFQKAETIPVLHEMI 83
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 CDLPQTHSLGSRRTLMLLAQMRRISLFSCLKDRHDFGFPQEEFGNQFQKAETIPVLHEMI 60

Qy 84 QQIFNLFSTKDSSAAWDETLLDKFYTELYQQLNDLEACVIQGVGVTETPLMKEDSILAVR 143
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61 QQIFNLFSTKDSSAAWDETLLDKFYTELYQQLNDLEACVIQGVGVTETPLMKEDSILAVR 120

Qy 144 KYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTNLQESLRSKE 188
||||||||||||||||||||||||||||||||||||||||||||
Db 121 KYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTNLQESLRSKE 165

RESULT 12
US-07-952-840-1
; Sequence 1, Application US/07952840
; Patent No. 5460956
; GENERAL INFORMATION:
; APPLICANT: Reichert, Paul
; APPLICANT: Hammond,Gerald S.
; APPLICANT: Le, Hung V.
; APPLICANT: Nagabhushan, Tattanahalli L.
; APPLICANT: Trotta, Paul P.
; TITLE OF INVENTION: Method for Preparing Interferon
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schering-Plough Corporation
; STREET: One Giralda Farms
; CITY: Madison
; STATE: New Jersey
; COUNTRY: USA
; ZIP: [illegible]
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: [illegible]
; OPERATING SYSTEM: [illegible]
; SOFTWARE: [illegible]
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/952,840
; FILING DATE: [illegible]
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA: [illegible]
; PRIOR APPLICATION DATA: [illegible]
; ATTORNEY/AGENT INFORMATION:
; NAME: Lunn, Paul G.
; REGISTRATION NUMBER: [illegible]
; REFERENCE/DOCKET NUMBER: [illegible]
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: [illegible]
; TELEFAX: [illegible]
; TELEX: [illegible]
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 165 amino acid residues
; TYPE: [illegible]
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-07-952-840-1

Query Match [illegible]; Score [illegible]; DB 1; Length 165;
Best Local Similarity [illegible]; Pred. No. [illegible];
Matches 165; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

Qy 24 [illegible] 83
[illegible]
Db 1 CDLPQTHSLGSRRTLMLLAQMRRISLFSCLKDRHDFGFPQEEFGNQFQKAETIPVLHEMI 60

Qy 84 [illegible] 143
[illegible]
Db 61 QQIFNLFSTKDSSAAWDETLLDKFYTELYQQLNDLEACVIQGVGVTETPLMKEDSILAVR 120

Qy 144 [illegible] 188
[illegible]
Db 121 KYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTNLQESLRSKE 165

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: January 17, 2003, 06:54:46 ; Search time 11.7167 Seconds
(without alignments)
318.960 Million cell updates/sec

Title: US-09-915-873-4
Perfect score: 960
Sequence: 1 MALTFALLVALLVLSCKSSC.....EIMKSPSLSTNLQEHLFSKE 198

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 120991 seqs, 19878514 residues

Total number of hits satisfying chosen parameters: 120991

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpa/PCT_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpa/US08_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubpa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/2/pubpa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubpa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubpa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/2/pubpa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	960	100.0	188	9	US-10-087-325-2	Sequence 2, Appli
2	957	99.7	189	9	US-09-915-873-2	Sequence 2, Appli
3	848	88.3	165	10	US-09-977-034-9	Sequence 9, Appli
4	776.5	80.9	189	9	US-09-908-193-2	Sequence 2, Appl
5	772.5	80.5	189	9	US-09-908-193-30	Sequence 30, Appl
6	766.5	79.8	189	10	US-09-919-497-73	Sequence 73, Appl
7	728.5	75.9	166	10	US-09-977-034-11	Sequence 11, Appl
8	727.5	75.8	170	10	US-09-977-034-16	Sequence 16, Appl
9	721.5	75.2	166	10	US-09-559-671A-80	Sequence 80, Appl
10	721.5	75.2	166	10	US-09-977-034-12	Sequence 12, Appl
11	717.5	74.7	166	10	US-09-559-671A-75	Sequence 75, Appl
12	703.5	73.3	166	10	US-09-559-671A-76	Sequence 76, Appl
13	703.5	73.3	166	10	US-09-559-671A-78	Sequence 78, Appl
14	698.5	72.8	166	10	US-09-977-034-2	Sequence 2, Appli
15	696.5	72.6	200	9	US-09-908-193-6	Sequence 6, Appli
16	694.5	72.3	166	10	US-09-559-671A-79	Sequence 79, Appl
17	694.5	72.3	166	10	US-09-977-034-10	Sequence 10, Appl
18	694.5	72.3	166	10	US-09-977-034-18	Sequence 18, Appl
19	692.5	72.1	166	10	US-09-977-034-19	Sequence 19, Appl
20	691.5	72.0	166	10	US-09-977-034-17A-84	Sequence 84, Appl
21	690.5	71.9	166	10	US-09-977-034-17	Sequence 17, Appl
22	690.5	71.9	166	10	US-09-977-034-13	Sequence 13, Appl
23	686.5	71.5	166	9	US-09-908-193-31	Sequence 31, Appl
24	685.5	71.4	166	10	US-09-559-671A-84	Sequence 84, Appl
25	685.5	71.4	166	10	US-09-977-034-8	Sequence 8, Appl
26	684.5	71.3	166	10	US-09-559-671A-77	Sequence 77, Appl
27	684.5	71.3	166	10	US-09-977-034-17A-82	Sequence 82, Appl
28	683.5	71.2	166	10	US-09-977-034-17A-81	Sequence 81, Appl
29	681.5	71.0	166	10	US-09-977-034-17A-85	Sequence 85, Appl
30	678.5	70.7	166	10	US-09-977-034-17A-83	Sequence 83, Appl
31	675.5	70.4	166	10	US-09-977-034-14	Sequence 14, Appl
32	668.5	69.6	166	10	US-09-977-034-4	Sequence 4, Appl
33	667.5	69.5	166	10	US-09-977-034-20	Sequence 20, Appl
34	661.5	68.9	166	10	US-09-977-034-21	Sequence 21, Appl
35	660.5	68.8	166	10	US-09-977-034-22	Sequence 22, Appl
36	659.5	68.7	166	10	US-09-977-034-23	Sequence 23, Appl
37	658.5	68.6	166	10	US-09-977-034-24	Sequence 24, Appl
38	657.5	68.5	166	10	US-09-977-034-25	Sequence 25, Appl
39	656.5	68.4	166	10	US-09-977-034-26	Sequence 26, Appl
40	655.5	68.3	166	10	US-09-977-034-27	Sequence 27, Appl
41	654.5	68.2	166	10	US-09-977-034-28	Sequence 28, Appl
42	653.5	68.1	166	10	US-09-977-034-29	Sequence 29, Appl
43	652.5	68.0	166	10	US-09-977-034-30	Sequence 30, Appl
44	651.5	67.9	166	10	US-09-977-034-31	Sequence 31, Appl
45	650.5	67.8	166	10	US-09-977-034-32	Sequence 32, Appl

ALIGNMENTS

RESULT 1
US-10-087-325-2
; Sequence 2, Application US/10087325
; Publication No. US20020127224A1
; GENERAL INFORMATION:
; APPLICANT: Escary, Jean-Louis
; TITLE OF INVENTION: DNA Encoding Two Novel Human Diseases Related Proteins
; FILE REFERENCE: 21705-0003
; CURRENT APPLICATION NUMBER: US/10/087,325
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 60/272,705
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ IDS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-087-325-2

Query Match 100.0%; Score 960; DB 9; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.6e-94;
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MALTFALLVALLVLSCKSSCSLDCCEDKQHQSCLPGESCCSFDCGHMCLKPVNTECNPH 60

Db 1 MALTFALLVALLVLSCKSSCSLDCCEDKQHQSCLPGESCCSFDCGHMCLKPVNTECNPH 60

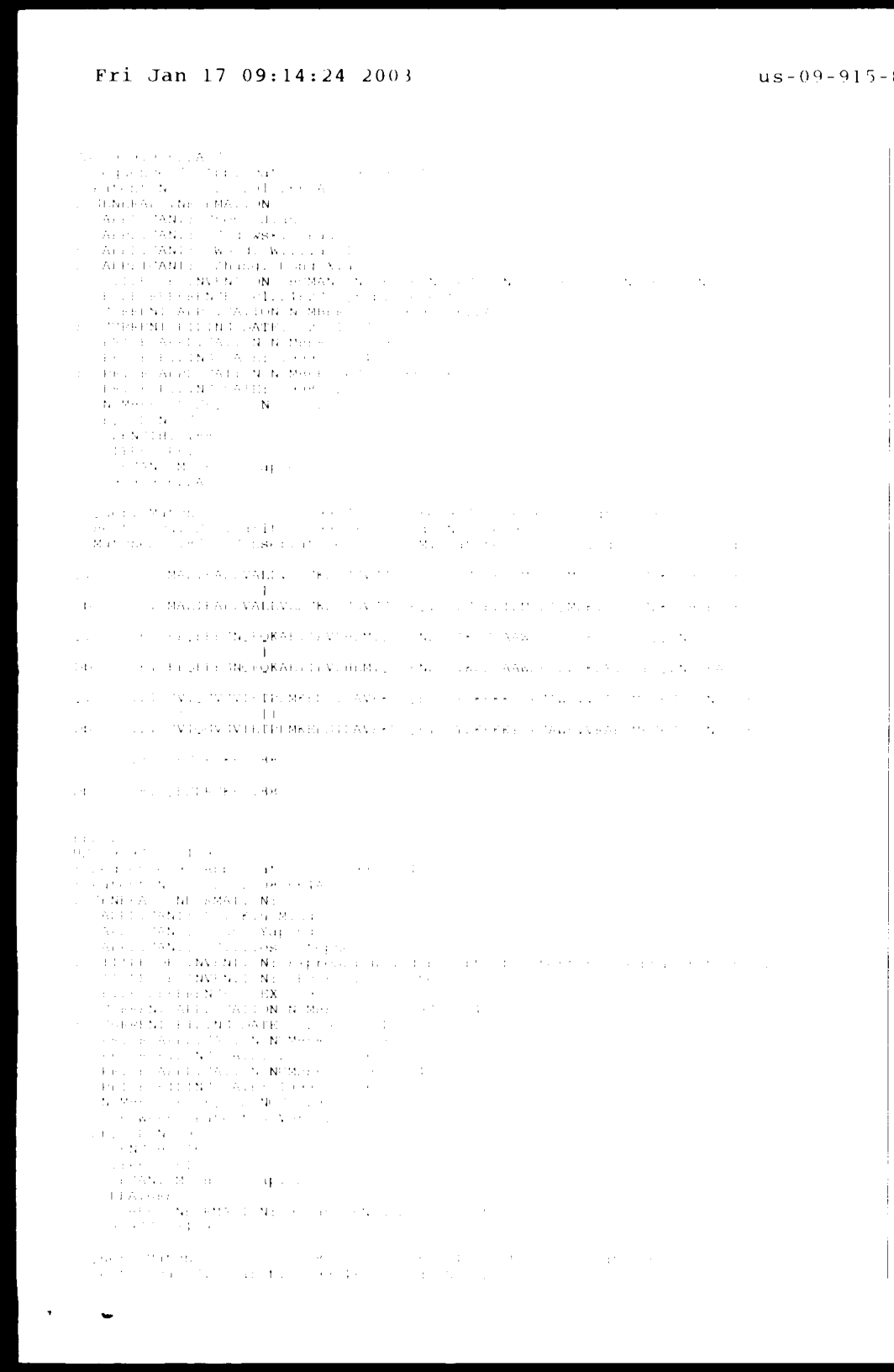
Qy 61 CVICTVTVITPIPMKLISIAVFVLMSAIFCYSMYAMVVFAGLKPPKPIKDEEDDNLQIVN 120

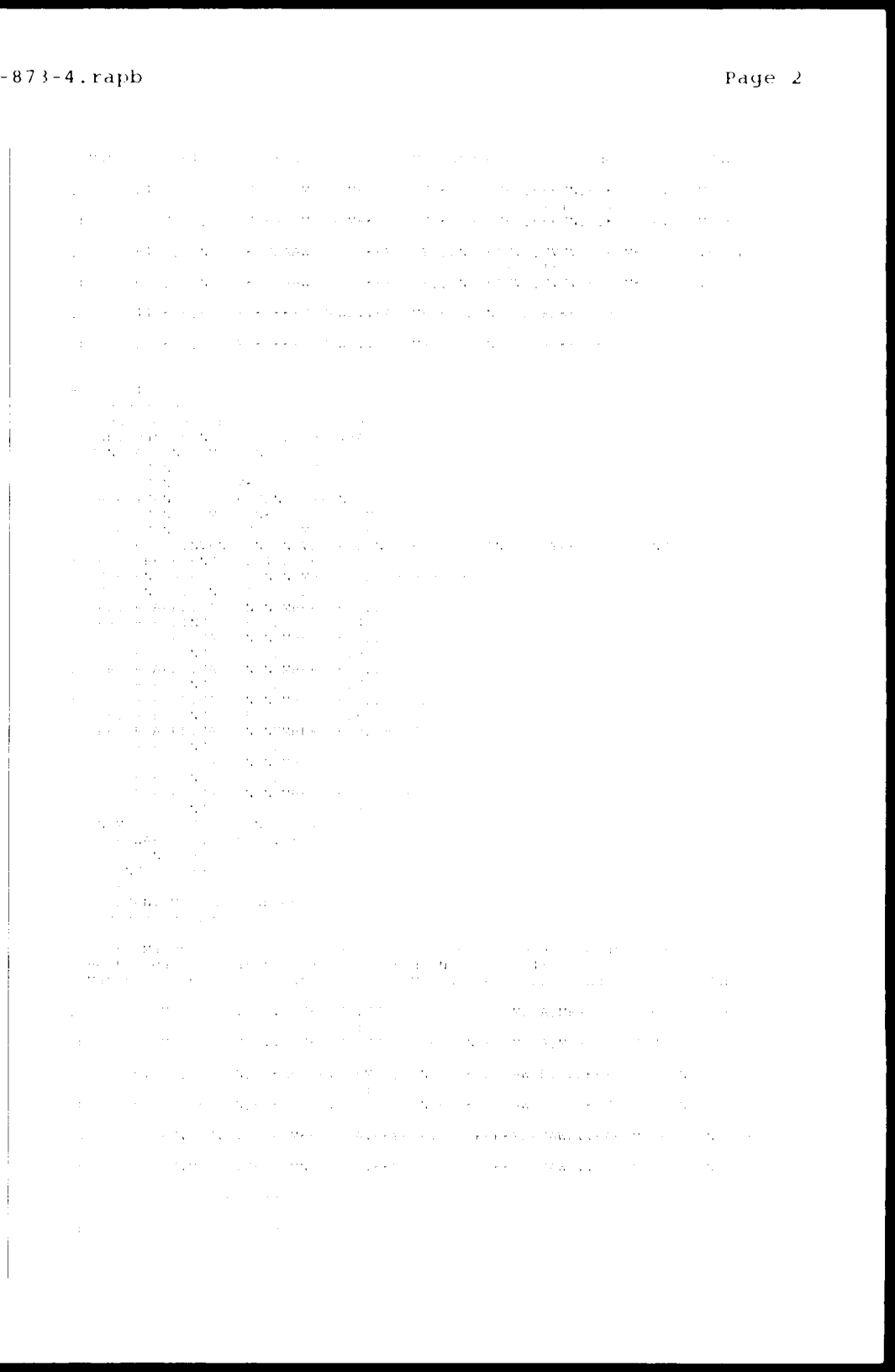
Db 61 CVICTVTVITPIPMKLISIAVFVLMSAIFCYSMYAMVVFAGLKPPKPIKDEEDDNLQIVN 120

Qy 121 EQSLSTNLQEHLFSKE 188

Db 121 EQSLSTNLQEHLFSKE 188

RESULT 2





US-09-908-193-30
; Sequence 30, Application US/09908193
; Publication No. US20020192748A1
; GENERAL INFORMATION:
; APPLICANT: RASTELLI, LUCA
; APPLICANT: SHIMKETS, RICHARD A.
; APPLICANT: ZERHUSEN, BRYAN
; APPLICANT: MALYANKAR, URIEL M.
; APPLICANT: PADIGARU, MURALIDHARA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND POLYPEPTIDES ENCODED THEREBY
; FILE REFERENCE: 21402-062
; CURRENT APPLICATION NUMBER: US/09/908,193
; CURRENT FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: 60/220,273
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: 60/221,650
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 60/221,333
; PRIOR FILING DATE: 2000-07-27
; PRIOR APPLICATION NUMBER: 60/220,912
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/218,875
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/218,870
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/218,901
; PRIOR FILING DATE: 2000-07-18
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 30
; LENGTH: 189
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-908-193-30

Query Match 80.5%; Score 772.5; DB 9; Length 189;
Best Local Similarity 82.0%; Pred. No. 4e-74;
Matches 155; Conservative 10; Mismatches 23; Indels 1; Gaps 1;

Qy 1 MALTPALIVALLVLSCKSSCSLGCDLPQTHSLGSRRTLMLLAQMSRISPSSCLMDRHDFG 60
|| |||| |:|||||||||:|||||:|||| :|||||||||| |:| |:| ||:|:|
Db 1 MASPFALLMVLVVLSCKSSCSLGCDLPETHSLDNRRTLMLLAQMSRISPSSCLMDRHDFG 60

Qy 61 FPQEFFDGNQFQKAETIVLEEIIQQIFNLFTKDSSAAWDEDLLDKFCTELYQQLNDLE 119
|||||| ||||||| | ||||:|||||||:|||||||||| |||| ||.|||:|:|
Db 61 FPQEFFDGNQFQKAPAISVLHELIQQIFNLFTKDSSAAWDEDLLDKFCTELYQQLNDLE 120

Qy 120 ACVIQEVGVTETPLMNEDSILAVRKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTN 179
|||:| | ||||| |||||||:|||:|||||| ||||||||||||:|||||| ||||
Db 121 ACVMQEERVGETPLMNADSILAVRKYFRRITLYLTEKKYSPCAWEVVRAEIMRSLSFST- 179

Qy 180 LQESLRSKE 188
||| || ||
Db 181 LQERLRRKE 189

RESULT 6
US-09-919-497-73
; Sequence 73, Application US/09919497
; Patent No. US20020106662A1
; GENERAL INFORMATION:
; APPLICANT: Mutter, George L.
; TITLE OF INVENTION: PROGNOSTIC CLASSIFICATION OF ENDOMETRIAL CANCER
; FILE REFERENCE: B0801/7225
; CURRENT APPLICATION NUMBER: US/09/919,497
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/221,735
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 73
; LENGTH: 189
; TYPE: PRT

; ORGANISM: Homo sapiens
US-09-919-497-73

Query Match [illegible]
Best Local Similarity [illegible]
Matches [illegible]

Qy [illegible]
Db [illegible]

Qy [illegible]
Db [illegible]

Qy [illegible]
Db [illegible]

Qy 180 LQESLRSKE 188
Db 181 LQERLRRKE 189

RESULT 7
US-09-927-041-11
; Sequence 11, Application US/[illegible]
; Patent No. US20020106662A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Min Min
; APPLICANT: Sun, Yuping
; APPLICANT: [illegible], Stephen D.
; TITLE OF INVENTION: Expression and Diagnostic Uses of Interferon Alpha-Inducible Genes
; TITLE OF INVENTION: [illegible]
; FILE REFERENCE: LEX-0[illegible]
; CURRENT APPLICATION NUMBER: US/09/927,041
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: US 60/224,[illegible]
; PRIOR FILING DATE: 2000-08-[illegible]
; PRIOR APPLICATION NUMBER: US 60/[illegible]
; PRIOR FILING DATE: [illegible]
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 276
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Human IFN-alpha [illegible]
US-09-927-041-11

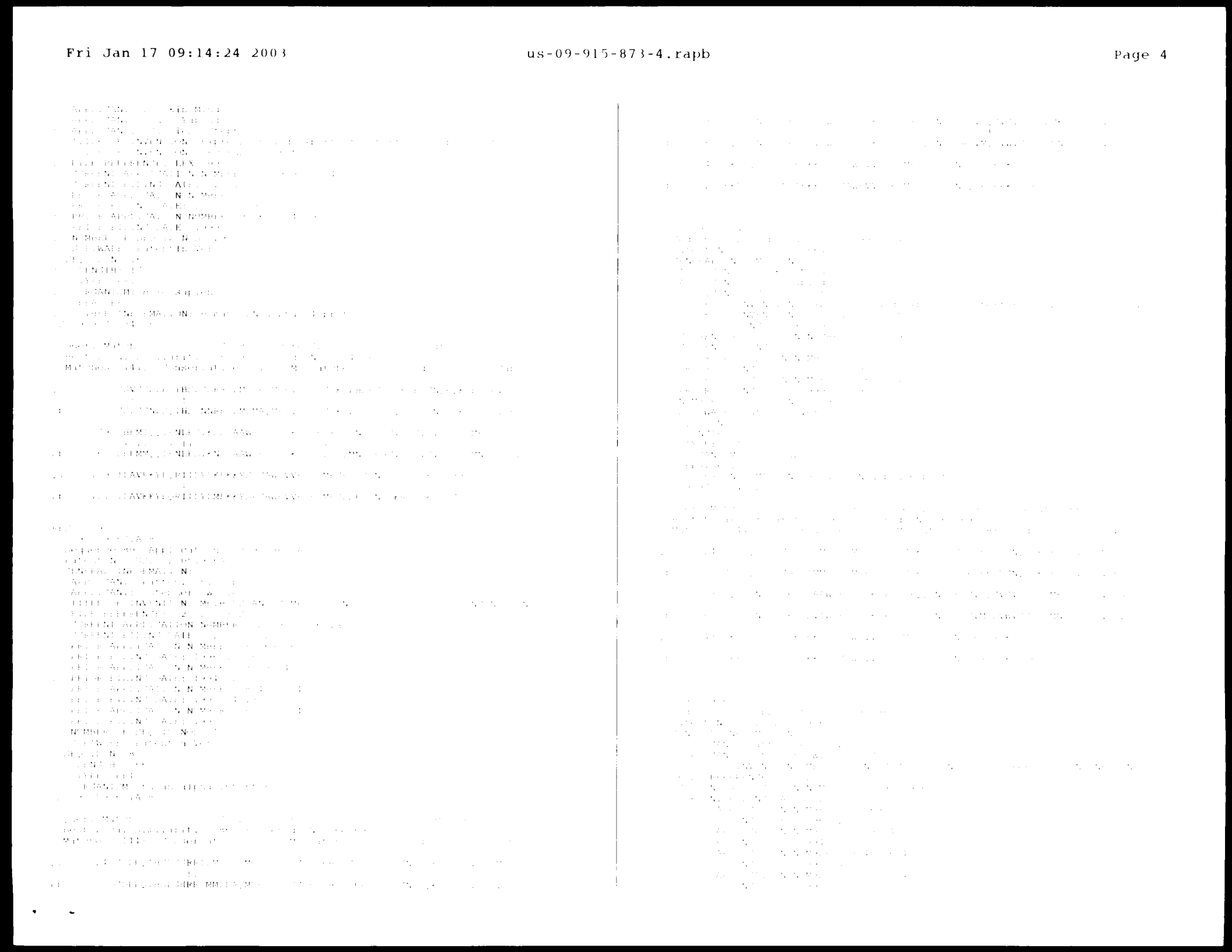
Query Match [illegible]
Best Local Similarity 85.5%; Pred. No. [illegible]
Matches [illegible]

Qy [illegible]
Db [illegible]

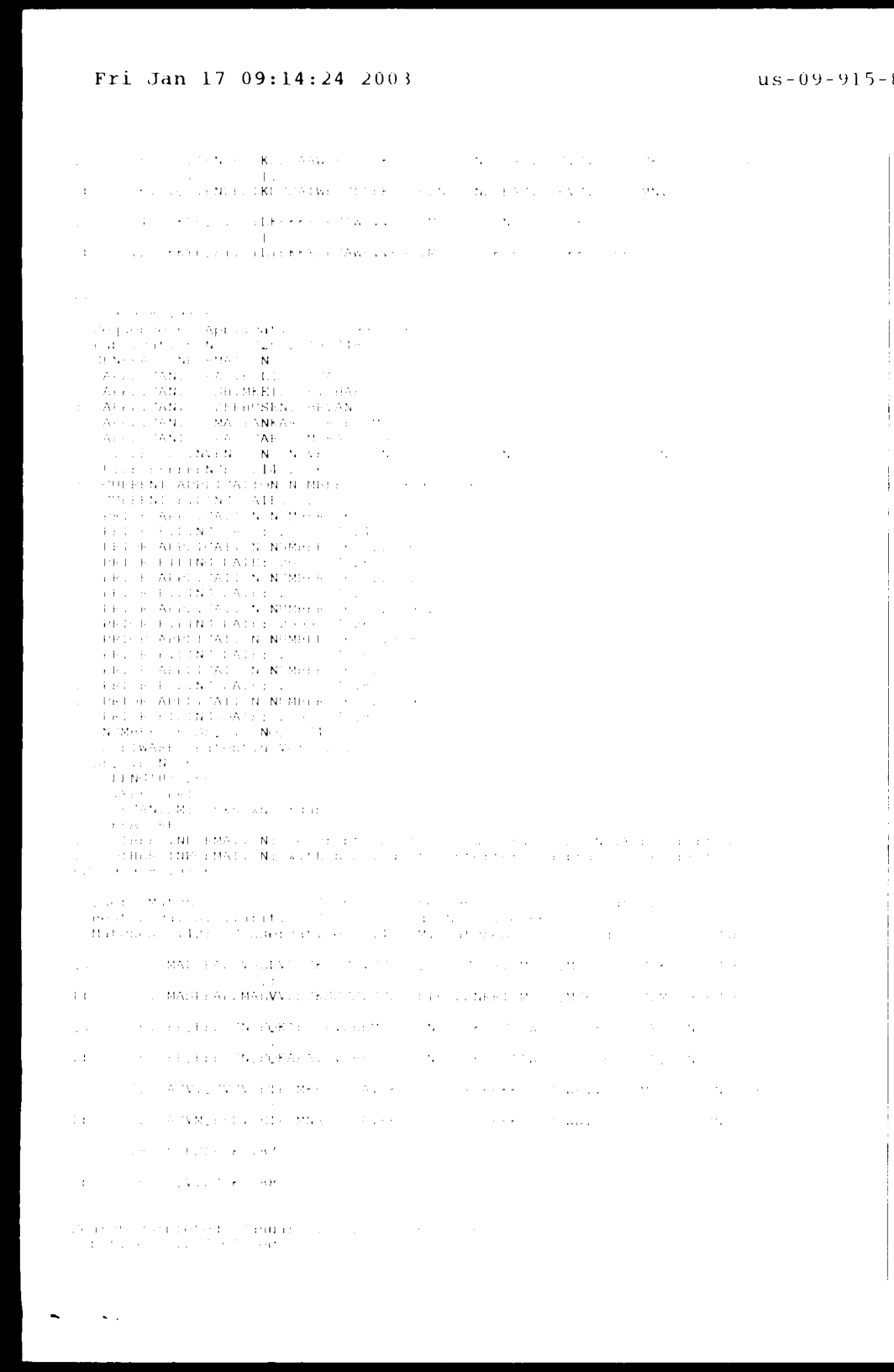
Qy [illegible]
Db [illegible]

Qy 143 EKYTQRTLHYLKEKKYSPCAWEVVRAEIMRSFSLSTNLQESLRSKE 188
Db 121 EKYTQRTLHYLKEKKYSPCAWEVVRAEIMRSLSANLQERLRRKE 166

RESULT 8
US-09-927-041-16
; Sequence 16, Application US/[illegible]
; Patent No. US20[illegible]
; GENERAL INFORMATION:



```
;  NUMBER OF SEQ ID NOS: 101
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 75
;   LENGTH: 166
;   TYPE: PRT
;   ORGANISM: consensus alpha interferon
US-09-559-671A-75

  Query Match             74.7%;  Score 717.5;  DB 1;  Length 166;
  Best Local Similarity   84.9%;  Pred. No. 1.6e-69;
  Matches  141;  Conservative    9;  Mismatches   15;  Indels    1;  Gaps    1;

Qy      24 CDLPQTHSLGSRRTLMLLAQMRRISLFSCLKDRHDFGFPQEEFGNQFQKAQAISVLHEM 83
           ||||||||||:|| |:|||||  ||| |||||||||||||||| |||||| : |||||
Db       1 CDLPQTHSLGNRRALILLAQMGRISPFSCLKDRHDFGFPQEEFDNQFQKAQAISVLHEM 60

Qy      84 IQQTFNLFSTKDSSAAWDETLLDKFYTELYQQLNDLEACVIQGVGVTETPLMKEDSILAV 143
           ||| |||||||||||:::|:|| ||||||||||||||| ||| |||| ||||||
Db      61 IQQTFNLFSTEDSSAAWDQSLLEKFSIELYQQLNDLEACVIQEVGVEETPLMNEDSILAV 120

Qy     144 RKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTNLQESLRSKE 189
           |||||||||||  |||||||||||||||||||| | ||||| || ||
Db     121 RKYFQRITLYLTEKKYSPCAWEVVRAEIMRSLSLSTNMQERLRRKD 166


RESULT 12
US-09-559-671A-76
; Sequence 76, Application US/09559671A
; Patent No. US20020051976A1
; GENERAL INFORMATION:
;  APPLICANT: Patton, Phillip
;  APPLICANT:  Stemmer, Willem P.C.
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR POLYPEPTIDE ENGINEERING
;  FILE REFERENCE: 02-020503US
;  CURRENT APPLICATION NUMBER: US/09/559,671A
;  CURRENT FILING DATE:  2001-05-25
;  PRIOR APPLICATION NUMBER: 08/709,041
;  PRIOR FILING DATE: 1996-12-18
;  PRIOR APPLICATION NUMBER: 08/198,431
;  PRIOR FILING DATE: 1994-02-17
;  PRIOR APPLICATION NUMBER: 08/125,694
;  PRIOR FILING DATE: 1995-04-18
;  PRIOR APPLICATION NUMBER: 08/547,874
;  PRIOR FILING DATE: 1995-10-30
;  NUMBER OF SEQ ID NOS: 101
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 76
;   LENGTH: 166
;   TYPE: PRT
;   ORGANISM: human alpha interferon
US-09-559-671A-76

  Query Match             73.3%;  Score 703.5;  DB 1;  Length 166;
  Best Local Similarity   82.5%;  Pred. No. 4.8e-68;
  Matches  137;  Conservative   12;  Mismatches   16;  Indels    1;  Gaps    1;

Qy      24 CDLPQTHSLGSRRTIMLLAQMRRISLFSCLKDRHDFGFPQEEFGNQFQKAQAISVLHEM 83
           ||||||||||:|| |:|||||  ||| |||||||||||||||| ||||| : |||||
Db       1 CDLPQTHSLGNRRALIILAQMGRISPFSCLKDRHDFGLPQEEFDNQFQKTQAISVLHEM 60

Qy      84 IQQIFNLFSTKDSSAAWDETLLDKFYTELYQQLNDLEACVIQGVGVTETPLMKEDSILAV 143
           ||| |||||:|||||:::|:|| |||||||:|||||| || |||| :|||||
Db      61 IQQTFNIFSTEDSSAAWDQSLLEKFSTELYQQLNNLEACVIQEVGMEETPLMNEDSILAV 120

Qy     144 RKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTNLQESLRSKE 189
           |||||||||||  |||||||||||:|||||| | ||||| || ||
Db     121 RKYFQRITLYLTEKKYSPCAWEVVRAEIMRSLSFSTNIQKRLRRKD 166


RESULT 13
US-09-559-671A-78
; Sequence 78, Application US/09559671A
; Patent No. US20020051976A1
; GENERAL INFORMATION:
;  APPLICANT: Patton, Phillip
;  APPLICANT:  Stemmer, Willem P.C.
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR POLYPEPTIDE ENGINEERING
;  FILE REFERENCE: 02-020503US
;  CURRENT APPLICATION NUMBER: US/09/559,671A
;  CURRENT FILING DATE:  2001-05-25
;  PRIOR APPLICATION NUMBER: 08/709,041
;  PRIOR FILING DATE: 1996-12-18
;  PRIOR APPLICATION NUMBER: 08/198,431
;  PRIOR FILING DATE: 1994-02-17
;  PRIOR APPLICATION NUMBER: 08/125,694
;  PRIOR FILING DATE: 1995-04-18
;  PRIOR APPLICATION NUMBER: 08/547,874
;  PRIOR FILING DATE: 1995-10-30
;  NUMBER OF SEQ ID NOS: 101
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 78
;   LENGTH: 166
;   TYPE: PRT
;   ORGANISM: human alpha interferon
US-09-559-671A-78

  Query Match             73.3%;  Score 703.5;  DB 1;  Length 166;
  Best Local Similarity   82.5%;  Pred. No. 4.8e-68;
  Matches  137;  Conservative   12;  Mismatches   16;  Indels    1;  Gaps    1;

Qy      24 CDLPQTHSLGSRRTIMLLAQMRRISLFSCLKDRHDFGFPQEEFGNQFQKAQAISVLHEM 83
           ||||||||||:|| |:|||||  ||| |||||||||||||||| ||||| : |||||
Db       1 CDLPQTHSLGNRRALIILAQMGRISPFSCLKDRHDFGLPQEEFDNQFQKTQAISVLHEM 60

Qy      84 IQQIFNLFSTKDSSAAWDETLLDKFYTELYQQLNDLEACVIQGVGVTETPLMKEDSILAV 143
           ||| |||||:|||||:::|:|| |||||||:|||||| || |||| :|||||
Db      61 IQQTFNIFSTEDSSAAWDQSLLEKFSTELYQQLNNLEACVIQEVGMEETPLMNEDSILAV 120

Qy     144 RKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTNLQESLRSKE 189
           |||||||||||  |||||||||||:|||||| | ||||| || ||
Db     121 RKYFQRITLYLTEKKYSPCAWEVVRAEIMRSLSFSTNIQKRLRRKD 166


RESULT 14
US-09-559-671A-2
; Sequence 2, Application US/09559671A
; Patent No. US20020051976A1
; GENERAL INFORMATION:
;  APPLICANT: Lee, Kin Mun
;  APPLICANT:  Sun, Yuqiu
;  APPLICANT:  Gillies, Stephen D.
;  TITLE OF INVENTION: Expression and Export of Interferon Alpha-2b as a Fusion
;  TITLE OF INVENTION: Protein
;  FILE REFERENCE: IEX-002
;  CURRENT APPLICATION NUMBER: US/09/559,671A
;  CURRENT FILING DATE:  2001-05-25
;  PRIOR APPLICATION NUMBER: 60/130,427
;  PRIOR FILING DATE: 1999-04-21
;  PRIOR APPLICATION NUMBER: 60/111,237
;  PRIOR FILING DATE: 1998-12-07
;  NUMBER OF SEQ ID NOS: 29
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 166
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-559-671A-2

  Query Match             73.3%;  Score 703.5;  DB 2;  Length 166;
  Best Local Similarity   82.5%;  Pred. No. 4.8e-68;
  Matches  137;  Conservative   12;  Mismatches   16;  Indels    1;  Gaps    1;

Qy      24 CDLPQTHSLGSRRTIMLLAQMRRISLFSCLKDRHDFGFPQEEFGNQFQKAQAISVLHEM 83
           ||||||||||:|| |:|||||  ||| |||||||||||||||| ||||| : |||||
Db       1 CDLPQTHSLGNRRALIILAQMGRISPFSCLKDRHDFGLPQEEFDNQFQKTQAISVLHEM 60

Qy      84 IQQIFNLFSTKDSSAAWDETLLDKFYTELYQQLNDLEACVIQGVGVTETPLMKEDSILAV 143
           ||| |||||:|||||:::|:|| |||||||:|||||| || |||| :|||||
Db      61 IQQTFNIFSTEDSSAAWDQSLLEKFSTELYQQLNNLEACVIQEVGMEETPLMNEDSILAV 120

Qy     144 RKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTNLQESLRSKE 189
           |||||||||||  |||||||||||:|||||| | ||||| || ||
Db     121 RKYFQRITLYLTEKKYSPCAWEVVRAEIMRSLSFSTNIQKRLRRKD 166
```



GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: January 17, 2003, 05:22:27 ; Search time 15.4774 Seconds
(without alignments)
479.674 Million cell updates/sec

Title: US-09-915-873-5
Perfect score: 851
Sequence: 1 CDLPQTHSLGSRRTLMLLAQ.....EIMRSFSLSTNLQESLRSKE 165

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseq-[illegible]:*
2: /SIDS2/gcgdata/geneseq/geneseq-[illegible]:*
3: /SIDS2/gcgdata/geneseq/geneseq-[illegible]:*
4: /SIDS2/gcgdata/geneseq/geneseq-[illegible]:*
5: /SIDS2/gcgdata/geneseq/geneseq-[illegible]:*
6: /SIDS2/gcgdata/geneseq/geneseq-[illegible]:*
7: /SIDS2/gcgdata/geneseq/geneseq-[illegible]AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseq-[illegible]:*
9: /SIDS2/gcgdata/geneseq/geneseq-[illegible]AA1999.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseq-[illegible]AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseq-[illegible]:*
12: /SIDS2/gcgdata/geneseq/geneseq-[illegible]:*
13: /SIDS2/gcgdata/geneseq/geneseq-[illegible]:*
14: /SIDS2/gcgdata/geneseq/geneseq-[illegible]:*
15: /SIDS2/gcgdata/geneseq/geneseq-[illegible]:*
16: /SIDS2/gcgdata/geneseq/geneseq-[illegible]:*
17: /SIDS2/gcgdata/geneseq/geneseq-[illegible]:*
18: /SIDS2/gcgdata/geneseq/geneseq-[illegible]:*
19: /SIDS2/gcgdata/geneseq/geneseq-[illegible]:*
20: /SIDS2/gcgdata/geneseq/geneseq-[illegible]:*
21: /SIDS2/gcgdata/geneseq/geneseq-[illegible]:*
22: /SIDS2/gcgdata/geneseq/geneseq-[illegible]:*
23: /SIDS2/gcgdata/geneseq/geneseq-[illegible]:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	851	100.0	165	6	AAP50228	Interferon alpha 2
2	851	100.0	165	15	AAR62234	Recombinant human
3	851	100.0	165	18	AAW43379	Human interferon a
4	851	100.0	165	18	AAW14015	Interferon-alpha 2
5	851	100.0	165	20	AAY43488	A human interferon
6	851	100.0	165	20	AAY43489	A human interferon
7	851	100.0	165	20	AAY43490	A human interferon
8	851	100.0	165	20	AAY02090	A multifunctional
9	851	100.0	165	20	AAY02091	A multifunctional
10	851	100.0	165	21	AAY82883	Interferon alpha 2

ALIGNMENTS

RESULT 1
[illegible]
ID AAP50228 standard; protein; [illegible] AA.
XX
AC AAP50228;
XX
DT [illegible] (first entry)
XX
DE Interferon alpha 2.
XX
KW Interferon alpha 2; hybrid protein; [illegible].
XX
FH Key Location/Qualifiers
FT Region 92..98
FT /label= [illegible] A
FT Region 87..91
FT /label= [illegible] B
FT Region 87..[illegible]
FT /label= [illegible]
FT Region 87..[illegible]
FT /label= [illegible]
FT Region 100..112
FT /label= [illegible]
FT Region 115..119
FT /label= [illegible]
FT Region 125..[illegible]
FT /label= [illegible]
FT Region 15..[illegible]
FT /label= [illegible] B
FT Region 15..[illegible]
FT /label= [illegible]
FT Region 87..[illegible]
FT /label= [illegible]

CC selected cells. The complex comprises the gene encoding interferon
CC releasably linked to a conjugate of nucleic acid binding agent and a
CC ligand which binds to a component on the surface of the cell. The method
CC is used for targeted expression of recombinant IFN in selected cells, in
CC vivo or in vitro, particularly for treatment of hepatitis, several forms
CC of cancer and leukaemia and condyloma acuminatum, or for production of
CC IFN for subsequent administration as exogenous protein.
XX
SQ Sequence 165 AA;

Query Match 100.0%; Score 851; DB 18; Length 165;
Best Local Similarity 100.0%; Pred. No. 5.9e-81;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CDLPQTHSLGSRRTLMLLAQMRRISLFSCLKDRHDFGFPQEEFGNQFQKAETIPVLHEMI 60
Db 1 CDLPQTHSLGSRRTLMLLAQMRRISLFSCLKDRHDFGFPQEEFGNQFQKAETIPVLHEMI 60

Qy 61 QQIFNLFSTKDSSAAWDETLLDKFYTELYQQLNDLEACVIQGVGVTETPLMKEDSILAVR 120
Db 61 QQIFNLFSTKDSSAAWDETLLDKFYTELYQQLNDLEACVIQGVGVTETPLMKEDSILAVR 120

Qy 121 KYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTNLQESLRSKE 165
Db 121 KYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTNLQESLRSKE 165

RESULT 4
AAW14015
ID AAW14015 standard; protein; 165 AA.
XX
AC AAW14015;
XX
DT 27-MAY-1997 (first entry)
XX
DE Interferon-alpha-2b.
XX
KW Interferon alpha-2b; recombinant; human; crystalline metal IFN alpha;
KW controlled release system, IFNalpha2b.
XX
OS Homo sapiens.
XX
PN US5602232-A.
XX
PD 11-FEB-1997.
XX
PF 25-FEB-1993; 93US-0024390.
XX
PR 14-DEC-1994; 94US-0356021.
PR 25-FEB-1993; 93US-0024390.
XX
PA (SCHE) SCHERING CORP.
XX
PI Hruza A, Menemar C, Nagabhushan N, Nagabhushan TL;
PI Reichert P, Tindall S;
XX
DR WPI; 1997-131840/12.
XX
PT Prodn. of crystalline metal interferon alpha-2 - from interferon and
PT metal acetate
XX
PS Disclosure; Column 13-14; 10pp; English.
XX
CC This sequence represents the recombinant human interferon alpha-2b
CC (IFNalpha2b). A crystalline metal version of this sequence is produced
CC by the method of the invention. The method of the invention comprises
CC forming an aqueous solution of IFNalpha2 and a metal acetate salt (either
CC zinc or cobalt acetate) at a pH of from 5 to 7, where the solution has
CC an initial concentration of from 5-80 mg/ml IFNalpha2 and 70-120 mM metal
CC acetate salt. The solution is then warmed from an initial temperature of
CC about 4 degrees C to 22 degrees C, until supersaturation occurs and metal
CC IFNalpha2 crystals appear. The method is used to produce crystalline
CC IFNalpha2 which is pharmaceutically useful in controlled release

CC IFNalpha2 formulations. The [illegible] proteins have been [illegible]
XX
SQ Sequence 165 AA;

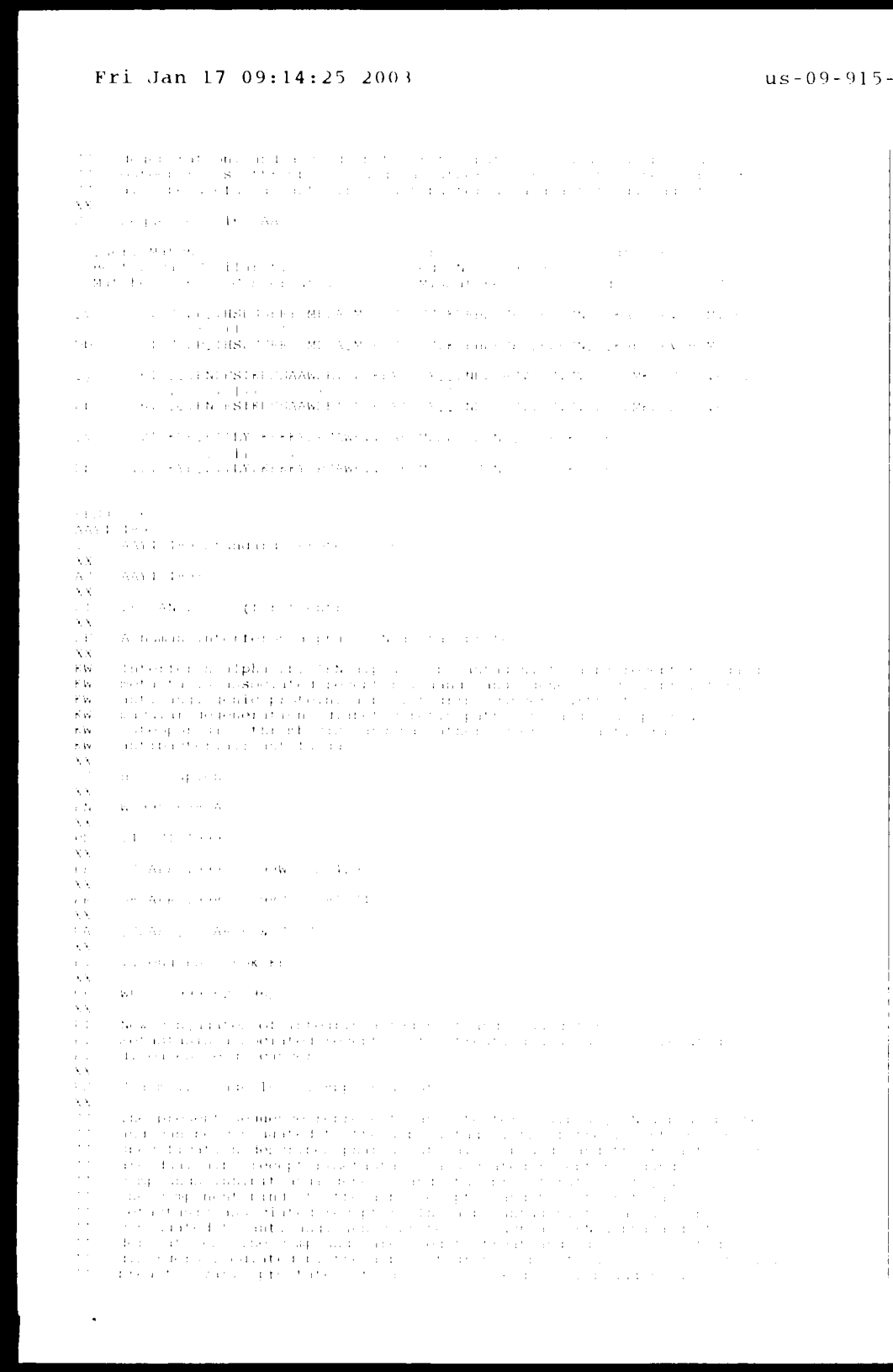
Query Match [illegible]
Best Local Similarity [illegible]
Matches [illegible]

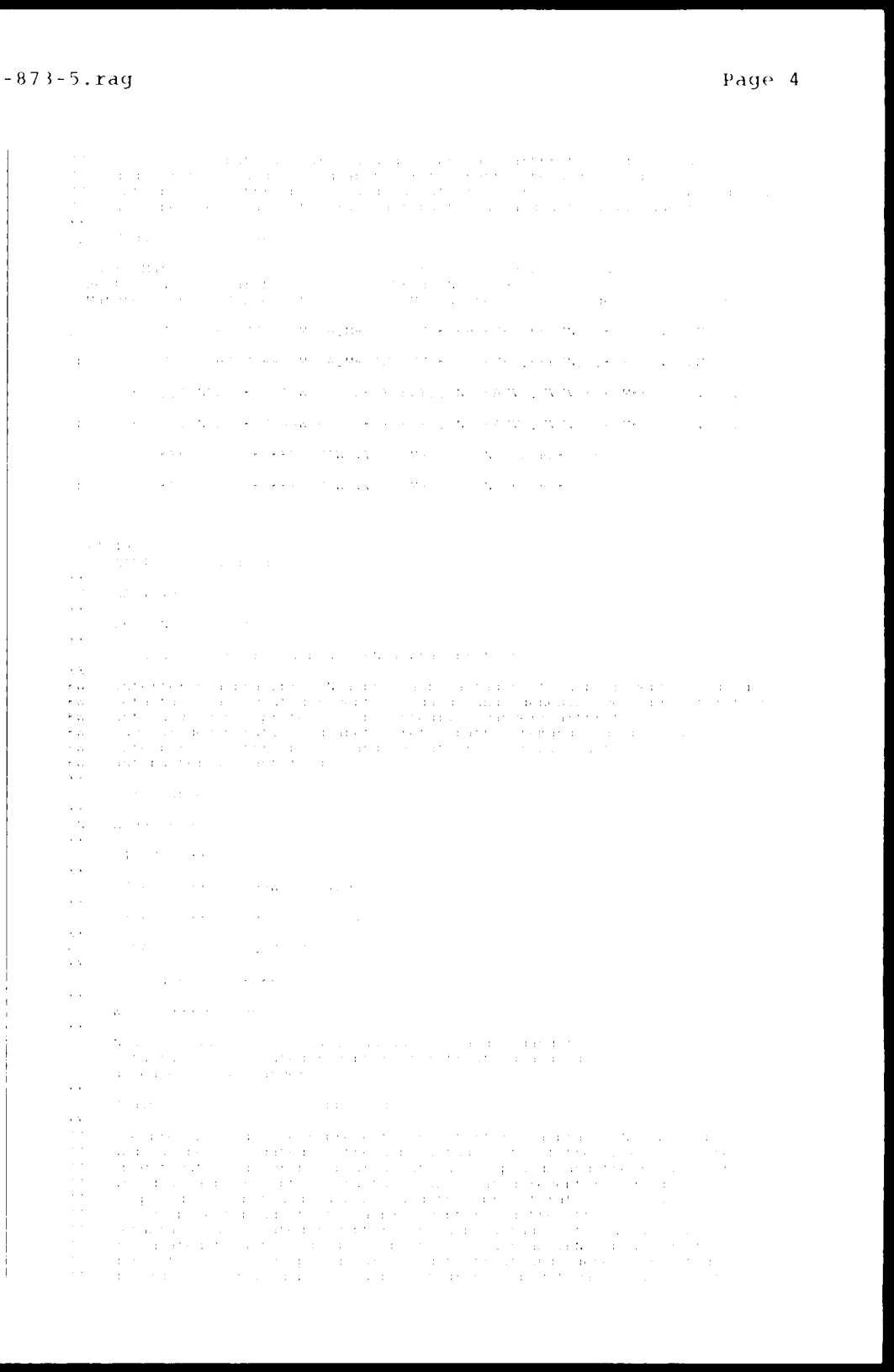
Qy [illegible]
Db [illegible]

Qy [illegible]
Db [illegible]

Qy [illegible]
Db [illegible]

RESULT 5
AAY14308
ID AAY14308 standard; Protein; 165 AA.
XX
AC AAY14308;
XX
DT 26-JAN-2000 (first entry)
XX
DE A human interferon alpha (IFN alpha) protein.
XX
KW Interferon-alpha-2b; IFN alpha; anti-metastatic; anti-neoplastic;
KW metastasis-associated receptor; cirrhosis; antiosteoporotic;
KW anti-angiogenic protein; anti-leukemic; cancer; arthritis;
KW macular degeneration; diabetic retinopathy; hemangioma; psoriasis;
KW osteoporosis; thrombosis; anaemia; atherosclerosis; AIDS; HIV;
KW antibacterial; antiherpes.
XX
OS Homo sapiens.
XX
PN WO9958136-A1.
XX
PD 18-NOV-1999.
XX
PF 07-APR-1999; 99WO-US007597.
XX
PR 08-APR-1998; 98US-0057078.
XX
PA (GLAX) GLAXO WELLCOME INC.
XX
PI Tjoelker LW, Kuo KK;
XX
DR WPI; 2000-063795/05.
XX
PT New conjugates of interferon, useful for treating [illegible]
PT metastasis associated receptor [illegible]
PT diseases, e.g. cancer
XX
PS Claim 12; Page 56-[illegible]; 109pp; English.
XX
CC The present sequence represents a human interferon alpha (IFN alpha)
CC and can be conjugated to the anti-metastatic [illegible]
CC specification describes pharmaceutical compositions and [illegible]
CC and dual-acting receptor/metastasis associated receptor [illegible]
CC compounds inhibit endothelial cell [illegible] proliferation [illegible]
CC one component binds to the cell surface [illegible]
CC metastasis-associated receptor (i). The anti-angiogenic [illegible]
CC conjugated to anti-angiogenic proteins, such as IFN alpha [illegible]
CC derivatives. The compounds are used to treat [illegible]
CC disorders (mediated by the anti-[illegible] specific receptor [illegible]
CC breast, ovary, prostate, stomach, [illegible] kidney [illegible]
CC melanoma, hepatoma, sarcoma and [illegible]. [illegible]





CC breast, ovary, prostate, stomach, colon, kidney or bladder, also
CC melanoma, hepatoma, sarcoma and lymphoma), arthritis and macular
CC degeneration, and also diabetic retinopathy, hemangioma, psoriasis,
CC osteoporosis, thrombosis, angina, atherosclerosis etc. The compounds may
CC also be useful as antiviral, antibacterial and antifungal agents.
XX
SQ Sequence 165 AA;

Query Match 100.0%; Score 851; DB 20; Length 165;
Best Local Similarity 100.0%; Pred. No. 5.9e-81;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CDLPQTHSLGSRRTLMLLAQMRRISLFSCLKDRHDFGFPQEEFDGNQFQKAEAISVLHEMI 60
Db 1 CDLPQTHSLGSRRTLMLLAQMRRISLFSCLKDRHDFGFPQEEFDGNQFQKAEAISVLHEMI 60

Qy 61 QQIFNLFSTKDSSAAWDETLLDKFYIELYQQLNDLEACVIQGVGVTETPLMKEDSILAVR 120
Db 61 QQIFNLFSTKDSSAAWDETLLDKFYIELYQQLNDLEACVIQGVGVTETPLMKEDSILAVR 120

Qy 121 KYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTNLQESLRSKE 165
Db 121 KYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTNLQESLRSKE 165

RESULT 8
AAY02090
ID AAY02090 standard; Protein; 165 AA.
XX
AC AAY02090;
XX
DT 16-JUL-1999 (first entry)
XX
DE A multifunctional protein of the invention.
XX
KW Angiostatin; endostatin; interferon; thrombospondin;
KW interferon-inducible protein; platelet factor 4; anti angiogenic;
KW anti-tumor; multifunctional protein; angiogenic-mediated disease;
KW cancer; diabetic retinopathy; macular degeneration; arthritis;
KW tumor cell production
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9916889-A1.
XX
PD 08-APR-1999.
XX
PF 30-SEP-1998; 98WO-US20464.
XX
PR 01-OCT-1997; 97US-0060609.
XX
PA (SEAR) SEARLE & CO G D.
XX
PI Bolanowski MA, Caparon MH, Casperson GF, Crowley SA;
PI Klein BK, McKearn JP;
XX
DR WPI; 1999-255098/21.
XX
PT New multifunctional proteins useful for treating angiogenic mediated
PT diseases
XX
PS Claim 5; Page 87; 121pp; English.
XX
CC The specification describes multifunctional proteins which comprise
CC combinations of angiostatin, endostatin, interferon, thrombospondin,
CC interferon-inducible protein and platelet factor 4, and have
CC anti-angiogenic and/or anti-tumor activity. The multifunctional protein
CC may exhibit useful properties such as having similar or greater
CC biological activity when compared to a single factor or by having
CC improved half-life or decreased adverse side effects, or a combination
CC of these properties. The proteins can be used for treating an
CC angiogenic-mediated disease, e.g. cancer, diabetic retinopathy, macular
CC degeneration or arthritis. They can also be used for inhibiting the
CC production of tumor cells (characterised in that breast, ovary,
CC prostate, pancreatic, gastric, colon, renal, bladder cancer, melanoma,
CC hepatoma, sarcoma and lymphoma) in a patient and for inhibiting tumor
CC growth. The present sequence represents a multifunctional protein of the
CC invention.
XX
SQ Sequence 165 AA;

Query Match 100.0%; Score 851; DB 20; Length 165;
Best Local Similarity 100.0%; Pred. No. 5.9e-81;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CDLPQTHSLGSRRTLMLLAQMRRISLFSCLKDRHDFGFPQEEFDGNQFQKAEAISVLHEMI 60
Db 1 CDLPQTHSLGSRRTLMLLAQMRRISLFSCLKDRHDFGFPQEEFDGNQFQKAEAISVLHEMI 60

Qy 61 QQIFNLFSTKDSSAAWDETLLDKFYIELYQQLNDLEACVIQGVGVTETPLMKEDSILAVR 120
Db 61 QQIFNLFSTKDSSAAWDETLLDKFYIELYQQLNDLEACVIQGVGVTETPLMKEDSILAVR 120

Qy 121 KYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTNLQESLRSKE 165
Db 121 KYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTNLQESLRSKE 165

RESULT 9
AAY02091
ID AAY02091 standard; Protein; 165 AA.
XX
AC AAY02091;
XX
DT 16-JUL-1999 (first entry)
XX
DE A multifunctional protein of the invention.
XX
KW Angiostatin; endostatin; interferon; thrombospondin;
KW interferon-inducible protein; platelet factor 4; anti angiogenic;
KW anti-tumor; multifunctional protein; angiogenic-mediated disease;
KW cancer; diabetic retinopathy; macular degeneration; arthritis;
KW tumor cell production
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9916889-A1.
XX
PD 08-APR-1999.
XX
PF 30-SEP-1998; 98WO-US20464.
XX
PR 01-OCT-1997; 97US-0060609.
XX
PA (SEAR) SEARLE & CO G D.
XX
PI Bolanowski MA, Caparon MH, Casperson GF, Crowley SA;
PI Klein BK, McKearn JP;
XX
DR WPI; 1999-255098/21.
XX
PT New multifunctional proteins useful for treating angiogenic mediated
PT diseases
XX
PS Claim 5; Page 87; 121pp; English.
XX
CC The specification describes multifunctional proteins which comprise
CC combinations of angiostatin, endostatin, interferon, thrombospondin,
CC interferon-inducible protein and platelet factor 4, and have
CC anti-angiogenic and/or anti-tumor activity. The multifunctional protein
CC may exhibit useful properties such as having similar or greater
CC biological activity when compared to a single factor or by having
CC improved half-life or decreased adverse side effects, or a combination
CC of these properties. The proteins can be used for treating an

PI Fok KF, Ijoeng FS;
XX
DR WPI; 2000-205894/18.
DR N-PSDB; AAZ93033.
XX
PT New bioconjugates comprising an avb3 antagonist and a
PT metastatic-associated receptor ligand, useful for treating cancer and
PT other angiogenic diseases, or as antiviral, antifungal or antibacterial
PT agents -
XX
PS Claim 13; Page 109; 123pp; English.
XX
CC Bioconjugates comprising one or more a_vb_3 antagonist moieties
CC coupled to a peptide or polypeptide having anti-angiogenic properties
CC can be used for treating a human patient with an
CC angiogenesis-mediated disease, e.g. cancer, arthritis, or macular
CC degeneration. The a_vb_3 integrin is normally associated with
CC endothelial cells but can promote the formation of blood vessels
CC (angiogenesis) in tumours. The a_vb_3 integrin is also known to
CC play a role in tumour metastasis, neoplasia, osteoporosis,
CC Paget's disease, retinopathy, arthritis, periodontal disease,
CC psoriasis and smooth muscle cell migration. Interferon alpha is a
CC family of proteins which possess complex antiviral, antineoplastic
CC and immunomodulating activities. Interferon alpha is effective
CC against a variety of cancers including hairy cell leukaemia,
CC chronic myelogenous leukaemia, malignant melanoma and Karposi's
CC sarcoma. Multi functional bioconjugates comprising both a_vb_3
CC antagonists and interferon alpha 2b can exhibit greater biological
CC activity when compared to a single factor or having improved
CC half-life or decreased adverse side effects, or a combination of
CC these properties. They can be used for inhibiting elevated levels
CC of tumor antigens, inhibiting the proliferation of tumor cells and
CC inhibiting tumor growth. The bioconjugates can also be used for
CC treating e.g. osteoporosis, humoral hypercalcemia of malignancy,
CC Paget's disease, retinopathy including diabetic retinopathy,
CC arthritis, including rheumatoid arthritis, periodontal disease,
CC psoriasis, thrombosis, angina, atherosclerosis, smooth muscle cell
CC migration and restenosis in a mammal. They are also useful as
CC antiviral, antifungal and antibacterial agents.
XX
SQ Sequence 165 AA;

Query Match 100.0%; Score 851; DB 21; Length 165;
Best Local Similarity 100.0%; Pred. No. 5.6e-81;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CDLPQTHSLGSRRTLMLLAQMRRISLFSCLKDRHDFGFPQEEFGNQFQKAETIPVLHEMI 60
Db 1 CDLPQTHSLGSRRTLMLLAQMRRISLFSCLKDRHDFGFPQEEFGNQFQKAETIPVLHEMI 60

Qy 61 QQIFNLFSTKDSSAAWDETLLDKFYTELYQQLNDLEACVIQGVGVTETPLMKEDSILAVR 120
Db 61 QQIFNLFSTKDSSAAWDETLLDKFYTELYQQLNDLEACVIQGVGVIETPLMKEDSILAVR 120

Qy 121 KYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTNLQESLRSKE 165
Db 121 KYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTNLQESLRSKE 165

RESULT 12
AAY82885
ID AAY82885 standard; Protein; 165 AA.
XX
AC AAY82885;
XX
DT 25-JUL-2000 (first entry)
XX
DE N-terminal modified interferon alpha 2b.
XX
KW Bioconjugate; a_vb_3 integrin; interferon alpha; angiogenesis;
KW cancer; tumour; osteoporosis; Paget's disease; Karposi's sarcoma;
KW periodontal disease; metastasis; neoplasia; retinopathy; arthritis;
KW psoriasis; leukaemia; malignant melanoma; atherosclerosis;
KW smooth muscle cell migration; inhibitor; treatment; arthritis;
KW angina; thrombosis; restenosis; antiviral; antifungal;
KW antibacterial.
XX
OS Homo sapiens.
XX
PN WO200065143 A1.
XX
PD 24-FEB-2000.
XX
PF 07 APR 1999; 99WO-US07206.
XX
PR 13-AUG-1998; 98US-0096411P.
XX
PA (SEAB-) SEARLE & CO G D.
XX
PI Fok KF, Ijoeng FS;
XX
DR WPI; 2000-205894/18.
DR N-PSDB; AAZ93034.
XX
PT New bioconjugates comprising an avb3 antagonist and a
PT metastatic-associated receptor ligand, useful for treating cancer and
PT other angiogenic diseases, or as antiviral, antifungal or antibacterial
PT agents -
XX
PS Claim 13; Page 110; 123pp; English.
XX
CC Bioconjugates comprising one or more a_vb_3 antagonist moieties
CC coupled to a peptide or polypeptide having anti-angiogenic properties
CC can be used for treating a human patient with an
CC angiogenesis-mediated disease, e.g. cancer, arthritis, or macular
CC degeneration. The a_vb_3 integrin is normally associated with
CC endothelial cells but can promote the formation of blood vessels
CC (angiogenesis) in tumours. The a_vb_3 integrin is also known to
CC play a role in tumour metastasis, neoplasia, osteoporosis,
CC Paget's disease, retinopathy, arthritis, periodontal disease,
CC psoriasis and smooth muscle cell migration. Interferon alpha is a
CC family of proteins which possess complex antiviral, antineoplastic
CC and immunomodulating activities. Interferon alpha is effective
CC against a variety of cancers including hairy cell leukaemia,
CC chronic myelogenous leukaemia, malignant melanoma and Karposi's
CC sarcoma. Multi functional bioconjugates comprising both a_vb_3
CC antagonists and interferon alpha 2b can exhibit greater biological
CC activity when compared to a single factor or having improved
CC half-life or decreased adverse side effects, or a combination of
CC these properties. They can be used for inhibiting elevated levels
CC of tumor antigens, inhibiting the proliferation of tumor cells and
CC inhibiting tumor growth. The bioconjugates can also be used for
CC treating e.g. osteoporosis, humoral hypercalcemia of malignancy,
CC Paget's disease, retinopathy including diabetic retinopathy,
CC arthritis, including rheumatoid arthritis, periodontal disease,
CC psoriasis, thrombosis, angina, atherosclerosis, smooth muscle cell
CC migration and restenosis in a mammal. They are also useful as
CC antiviral, antifungal and antibacterial agents. This is a
CC modified version of interferon alpha 2b such that the amino
CC terminus of interferon alpha 2b is Cys-IFN2b as opposed
CC to Ala-Cys-IFN2b.
XX
SQ Sequence 165 AA;

Query Match 100.0%; Score 851; DB 21; Length 165;
Best Local Similarity 100.0%; Pred. No. 5.6e-81;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CDLPQTHSLGSRRTLMLLAQMRRISLFSCLKDRHDFGFPQEEFGNQFQKAETIPVLHEMI 60
Db 1 CDLPQTHSLGSRRTLMLLAQMRRISLFSCLKDRHDFGFPQEEFGNQFQKAETIPVLHEMI 60

Qy 61 QQIFNLFSTKDSSAAWDETLLDKFYTELYQQLNDLEACVIQGVGVTETPLMKEDSILAVR 120
Db 61 QQIFNLFSTKDSSAAWDETLLDKFYTELYQQLNDLEACVIQGVGVIETPLMKEDSILAVR 120

Qy 121 KYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTNLQESLRSKE 165
Db 121 KYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTNLQESLRSKE 165

```
Qy     121 KYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSINLQESLRSKE 165
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 KYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSINLQESLRSKE 165


RESULT 15
AAM50705
ID   AAM50705 standard; Protein; 165 AA.
XX
AC   AAM50705;
XX
DT   15-APR-2002  (first entry)
XX
DE   Human mature interferon-alpha2b R121K,Q125R mutant.
XX
KW   Interferon-alpha2b; IFN-alpha; human; virucide; antiviral; hepatic;
KW   hepatitis; infection; liver; gene therapy; mutant; protein.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 121
FT                   /note= "Arg in wild-type sequence"
FT   Misc-difference 125
FT                   /note= "Gln in wild-type sequence"
XX
PN   US6331525-B1.
XX
PD   18-DEC-2001.
XX
PF   23-AUG-1999;   99US-0379434.
XX
PR   17-MAR-1997;   97US-0819238.
PR   14-MAR-1996;   96US-0616023.
XX
PA   (IMMU-) IMMUNE RESPONSE CORP.
XX
PI   Chiou HC, Carlo DJ;
XX
DR   WPI; 2002-105276/14.
XX
PT   Targeted delivery of nucleic acids encoding interferons to the liver
PT   for the treatment of hepatitis infections -
XX
PS   Example 3; Column 23-26; 28pp; English.
XX
CC   The present sequence corresponds to a mutant human mature
CC   interferon alpha2b (IFN-alpha2b) in which residue Arg-121 of the
CC   wild-type sequence (see AAM50693) is replaced by Arg, and wild-type
CC   Gln-125 is replaced by Arg, the residues found in the corresponding
CC   positions of IFN-alpha1. The mutations were introduced into
CC   IFN-alpha2b DNA by site-directed mutagenesis. The virucide
CC   activity of the full-length mutant protein expressed in mouse cells
CC   (taking the activity of the wild-type protein as 1) was 3000, with
CC   an enhancement factor of 3000. The invention provides a method for
CC   the targeted delivery of nucleic acids encoding interferons,
CC   especially human IFN-alpha2b, to the liver for the treatment of
CC   hepatitis infections.
XX
SQ   Sequence   165 AA;

  Query Match             100.0%;  Score 851;  DB 23;  Length 165;
  Best Local Similarity   100.0%;  Pred. No. 5.9e-81;
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CDLPQTHSLGSRRTLMLLAQMRRISLFSCLKDRHDFGFPQEEFGNQFQKAETIPVLHEMI 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CDLPQTHSLGSRRTLMLLAQMRRISLFSCLKDRHDFGFPQEEFGNQFQKAETIPVLHEMI 60

Qy      61 QQIFNLFSTKDSSAAWDETLLEKFYIELYQQLNDLEACVIQGVGVTETPLMKEDSILAVR 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 QQIFNLFSTKDSSAAWDETLLEKFYIELYQQLNDLEACVIQGVGVTETPLMKEDSILAVR 120
```

```
Db     [illegible]

Qy     121 KYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSINLQESLRSKE 165
           [illegible]
Db     121 KYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSINLQESLRSKE 165


Search completed: January 17, 2003, [illegible]
Job time : 17.8071 secs
```

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: January 17, 2003, 06:50:48 ; Search time 22.9[illegible] Seconds
(without alignments)
692.560 Million cell updates/sec

Title: US-09-915-873-5
Perfect score: 851
Sequence: 1 CDLPQTHSLGSRRT[illegible]MLLAQ..........[illegible] 165

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	851	100.0	165	2	I78570	alpha 2 interferon
2	848	99.6	188	1	IVHUA2	interferon alpha 2
3	728.5	85.6	189	1	IVHUA7	interferon alpha 5
4	721.5	84.8	189	1	IVHUI6	interferon alpha I
5	719.5	84.5	167	2	D25843	interferon alpha [illegible]
6	705.5	82.9	189	2	I52347	interferon alpha M
7	703.5	82.7	189	2	I51570	interferon precurs
8	701.5	82.4	189	1	IVHU14	interferon alpha I
9	695.5	81.7	176	2	I56314	interferon-alpha -
10	694.5	81.6	189	1	IVHU4B	interferon alpha I
11	692.5	81.4	181	2	I56313	interferon alpha 2
12	692.5	81.4	189	2	I84464	interferon-alpha F
13	691.5	81.3	189	1	IVHU16	interferon alpha I
14	690.5	81.1	167	2	E25843	interferon alpha F
15	690.5	81.1	189	1	IVHUF	interferon alpha I
16	690.5	81.1	189	1	IVHUA5	interferon alpha 5
17	688.5	80.9	189	1	IVHUA1	interferon alpha-I
18	685.5	80.6	189	2	I37584	IFN alpha N protei
19	684.5	80.4	189	1	IVHUA9	interferon alpha I
20	684.5	80.4	189	1	IVHUI8	interferon alpha I
21	678.5	79.7	167	2	F25843	interferon alpha J
22	678.5	79.7	189	1	IVHUA0	interferon alpha-7
23	678.5	79.7	189	2	I53102	interferon-alpha J
24	665.5	78.2	162	2	C25843	interferon alpha-B
25	664.5	78.1	189	1	IVHUA4	interferon alpha 4
26	637.5	74.9	184	1	IVHOA4	interferon alpha I
27	633.5	74.4	184	1	IVHOA2	interferon alpha I
28	632.5	74.3	184	1	IVHOA3	interferon alpha I
29	632.5	74.3	184	1	IVHOA1	interferon alpha-I
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
31	[illegible]	[illegible]	[illegible]	[illegible]	IVMSA[illegible]	[illegible]
32	[illegible]	[illegible]	[illegible]	[illegible]	IVMSA[illegible]	[illegible]
33	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
34	[illegible]	62.[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
35	[illegible]	62.2	[illegible]	[illegible]	[illegible]	[illegible]
36	528.5	62.1	[illegible]	[illegible]	[illegible]	[illegible]
37	525.5	61.8	[illegible]	1	[illegible]	[illegible]
38	525.5	61.8	[illegible]	1	[illegible]	[illegible]
39	525.5	61.8	[illegible]	2	[illegible]	[illegible]
40	525.5	61.8	192	1	[illegible]	interferon alpha [illegible]
41	524.5	61.6	190	1	IVMSA[illegible]	interferon alpha [illegible]
42	524.5	61.6	190	2	[illegible]	interferon alpha [illegible]
43	[illegible]	61.5	190	2	[illegible]	interferon alpha [illegible]
44	511.5	60.1	189	1	IVMSA[illegible]	interferon alpha [illegible]
45	509.5	59.9	189	[illegible]	[illegible]	[illegible]

ALIGNMENTS

RESULT 1
[illegible]
alpha 2 interferon - human (fragment)
C;Species: Homo sapiens (man)
C;Date: [illegible] #sequence_revision [illegible] #text_change [illegible]
C;Accession: I7857[illegible]
R;Weber, H.; Weissmann, C.
Nucleic Acids Res. 11, [illegible]
A;Title: Formation of genes coding for hybrid proteins by [illegible]
A;Reference number: [illegible]; MUID:[illegible]; PMID:[illegible]
A;Accession: I78570
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-165 <RES>
A;Cross-references: GB:M2[illegible]; NID:[illegible]; PIDN:AAA52715.1; [illegible]
C;Genetics:
A;Gene: IFNA
C;Superfamily: interferon alpha

Query Match [illegible]; Score [illegible]; DB 2; Length [illegible];
Best Local Similarity [illegible]; Pred. No. [illegible];
Matches 165; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

```
Qy   1 CDLPQTHSLGSRRTLMLLAQMR[illegible]
       [illegible]
Db   1 CDLPQTHSLGSRRTLMLLAQMR[illegible]

Qy  [illegible] [illegible]NESIK[illegible]AAW[illegible]NFTA[illegible]VIHEIMK[illegible]
       [illegible]
Db  [illegible] [illegible]NESIK[illegible]AAW[illegible]NFTA[illegible]VIHEIMK[illegible]

Qy  121 KYFQRITLYLKEKKYSPCAWEVVRA[illegible]SINLQESLRSKE 165
        [illegible]
Db  121 KYFQRITLYLKEKKYSPCAWEVVRA[illegible]SINLQESLRSKE 165
```

RESULT 2
IVHUA2
interferon alpha-2 precursor (allele [illegible]) [validated] - human
N;Alternate names: IFN-alpha2; interferon alpha-A; interferon alpha [illegible]
C;Species: Homo sapiens (man)
C;Date: [illegible] #sequence_revision [illegible] Sep 1981 #text_change [illegible]
C;Accession: A[illegible]; [illegible]
R;Goeddel, D.V.; Yelverton, E.; Ullrich, A.; Heyneker, H.L.; M[illegible]
ss, M.; Familletti, P.C.; Pestka, S.
Nature 287, 411-416, [illegible]
A;Title: Human leukocyte interferon produced by Escherichia [illegible]
A;Reference number: A[illegible]; MUID:[illegible]; PMID:[illegible]
A;Accession: A[illegible]
A;Molecule type: DNA
A;Residues: 1-188 <[illegible]>
A;Cross-references: GB:V[illegible]; NID:[illegible]; PIDN:CAA2[illegible]

A;Molecule type: DNA
A;Residues: 1-189 <HEN>
A;Cross-references: EMBL:X02956; NID:g32659; PIDN:CAA26792.1; PID:g[illegible]
R;Goeddel, D.V.; Leung, D.W.; Dull, T.J.; Gross, M.; Lawn, R.M.; McCandliss, R.; Seeburg [illegible]
Nature 290, 20-26, 1981
A;Title: The structure of eight distinct cloned human leukocyte interferon cDNAs.
A;Reference number: A93249; MUID:81148795; PMID:6163083
A;Accession: A01833
A;Molecule type: mRNA
A;Residues: 57-189 <GOE>
A;Cross-references: GB:V00541; GB:J00210; NID:g32718; PIDN:CAA23802.1; PID:g[illegible]
A;Note: eight classes of interferon alpha clones were identified; this sequence is deri[illegible]
C;Genetics:
A;Gene: GDB:IFNA5
A;Cross-references: GDB:136362; OMIM:147565
A;Map position: 9p22-9p22
C;Superfamily: interferon alpha
C;Keywords: antiviral; cytokine; leukocyte
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-189/Product: interferon alpha-5 #status predicted <MAT>

```
  Query Match             85.6%;  Score 728.5;  DB 1;  Length 189;
  Best Local Similarity   85.5%;  Pred. No. 4.2e-60;
  Matches  142;  Conservative    7;  Mismatches   16;  Indels    1;  Gaps    1;

Qy          1 CDLPQTHSLGSRRTLMLLAQMRRISLFSCLKDRHDFGFPQEEF-GNQFQKAETIPVLHEM 59
              ||||||||| :|||||:||| ||| ||||||||:||||||||| |||||:: | |||.|
Db         24 CDLPQTHSLSNRRTLMIMAQMGRISPFSCLKDRHDFGFPQEEFDGNQFQKAQAISVLHEM 83

Qy         60 IQQIFNLFSTKDSSAAWDETLLDKFYTELYQQLNDLEACVIQGVGVTETPLMKEDSILAV 119
              ||| ||||||||||| ||||||||||||||||||||||||::| ||| :|||  |||| |
Db         84 IQQTFNLFSTKDSSATWDETLLDKFYTELYQQLNDLEACMMQEVGVEDTPLMNVDSILTV 143

Qy        120 RKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTNLQESLRSKE 165
              ||||||||||| |||||||||||||||||||||||| |||| || ||
Db        144 RKYFQRITLYLTEKKYSPCAWEVVRAEIMRSFSLSANLQERLRRKE 189
```

RESULT 4
IVHUI6
interferon alpha-I-6 precursor - human
N;Alternate names: HuIFN-alpha-I-6; LeIF K; type I interferon
C;Species: Homo sapiens (man)
C;Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 18-Jun-1999
C;Accession: A23753
R;Henco, K.; Brosius, J.; Fujisawa, A.; Fujisawa, J.I.; Haynes, J.R.; Hochstadt, J.; Kov[illegible]
J. Mol. Biol. 185, 227-260, 1985
A;Title: Structural relationship of human interferon alpha genes and pseudogenes.
A;Reference number: A92916; MUID:86072205; PMID:4057246
A;Accession: A23753
A;Molecule type: DNA
A;Residues: 1-189 <HEN>
A;Cross-references: GB:X02958; NID:g32662; PIDN:CAA26701.1; PID:g[illegible]
C;Genetics:
A;Gene: GDB:IFNA6
A;Cross-references: GDB:136363; OMIM:147566
A;Map position: 9p22-9p22
C;Superfamily: interferon alpha
C;Keywords: antiviral
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-189/Product: interferon alpha-I-6 #status predicted <MAT>
F;24-122,52-162/Disulfide bonds: #status predicted

```
  Query Match             84.8%;  Score 721.5;  DB 1;  Length 189;
  Best Local Similarity   86.7%;  Pred. No. 1.9e-59;
  Matches  144;  Conservative    3;  Mismatches   18;  Indels    1;  Gaps    1;

Qy          1 CDLPQTHSLGSRRTLMLLAQMRRISLFSCLKDRHDFGFPQEEF-GNQFQKAETIPVLHEM 59
              ||||||||| |||:||||||||||||||||||||| |||||| ||||| || | ||||:
Db         24 CDLPQTHSLGHRRTMMLLAQMRRISLFSCLKDRHDFRFPQEEFDGNQFQKAEAISVLHEV 83

Qy         60 IQQIFNLFSTKDSSAAWDETLLDKFYTELYQQLNDLEACVIQGVGVTETPLMKEDSILAV 119
              ||| |||||||||| |||| |||| ||||||||||||||||:| | |  ||| ||||||
```

Db 84 IQQIFNLFSTKDSSVAWDERLLDKFYTELYQQLNDLEACVMQEVWVGGTPLMNEDSILAV [illegible]

Qy 120 RKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTNLQESLRSKE 16[illegible]

Db 144 RKYF[illegible]

RESULT 5
[illegible]
Interferon alpha [illegible] - human
N;Alternate names: human leukocyte interferon (IFN)
C;Species: Homo sapiens (man)
C;Date: 16-Aug-1988 #sequence_revision 16-Aug-1988 #text_change [illegible]
C;Accession: [illegible]
[illegible]
A;Title: Anomalous behavior [illegible]
A;Reference number: [illegible]
A;Accession: [illegible]
A;Status: nucleic acid sequence [illegible]
A;Molecule type: mRNA
A;Residues: [illegible]
C;Superfamily: interferon alpha

Query Match 84.[illegible]; Score [illegible]; DB [illegible]; Length [illegible];
Best Local Similarity [illegible]; Pred. No. [illegible];
Matches 14[illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

Qy 1 CDLPQTHS[illegible]
Db 2 CDLPQTHSLSNRRTLMIMAQM[illegible]

Qy 60 IQQIFNLFSTKD[illegible]
Db 62 IQQIFNLFSTKDSSAA[illegible]

Qy 120 RKYFQRITLYLKEKKYS[illegible]
Db 122 RKYFQRITLYL[illegible]

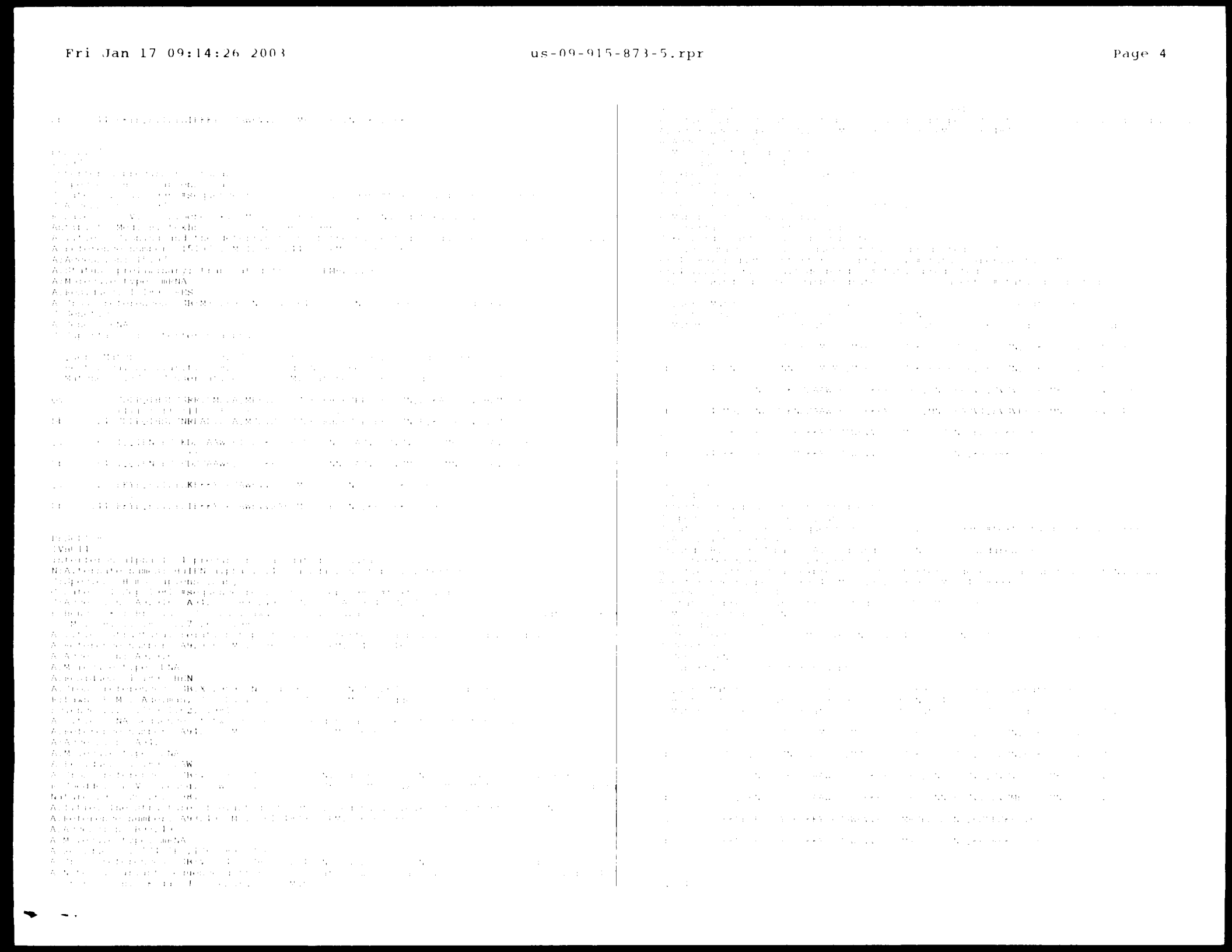
RESULT 6
I52347
interferon alpha M1 precursor - human
C;Species: Homo sapiens (man)
C;Date: 02-Jul-1996 #sequence_revision [illegible] 1996 #text_change [illegible]
C;Accession: I52347
R;Linnane, A.W.; Bellnarz, M.W.; McMullen, [illegible]
Biochem. Int. [illegible], 1984
A;Title: Nucleotide sequence [illegible]
A;Reference number: I52347; [illegible]
A;Accession: I52347
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-189 <RES>
A;Cross-references: GB:M[illegible]; NID:[illegible]; PIDN:AAA[illegible]
C;Genetics:
A;Gene: IFNA
C;Superfamily: interferon alpha

Query Match [illegible]
Best Local Similarity [illegible]; Pred. No. [illegible];
Matches 138; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

Qy 1 [illegible]
Db 24 CDLPQTHSL[illegible]

Qy 60 IQQIFNLFSTKDSSAAWDE[illegible]
Db 84 [illegible]FNLFST[illegible]DSSAAW[illegible]

Qy 120 RKYFQRITLYLKEKKYSPCAWEVVRAEIMRS[illegible]



interferon alpha-I-4b precursor - human
N;Alternate names: HuIFN-alpha-I-4b; type I interferon
C;Species: Homo sapiens (man)
C;Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 18-Jul-1997
C;Accession: E23753
R;Henco, K.; Brosius, J.; Fujisawa, A.; Fujisawa, J.I.; Haynes, J.R.; Hochstadt, J.; K...
J. Mol. Biol. 185, 227-260, 1985
A;Title: Structural relationship of human interferon alpha genes and pseudogenes.
A;Reference number: A92916; MUID:86037205; PMID:4057246
A;Accession: E23753
A;Molecule type: DNA
A;Residues: 1-189 <HEN>
A;Cross-references: GB:X02955; NID:g32656; PIDN:CAA26731.1; PID:g32657
C;Genetics:
A;Gene: GDB:IFN1
A;Cross-references: GDB:119328; OMIM:147660
A;Map position: 9p22-9p22
C;Superfamily: interferon alpha
C;Keywords: antiviral
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-189/Product: interferon alpha-I-4b #status predicted MAT
F;24-122/52 162/Disulfide bonds: #status predicted

Query Match 81.6%; Score 694.5; DB 1; Length 189;
Best Local Similarity 81.9%; Pred. No. 5.9e-57;
Matches 136; Conservative 12; Mismatches 17; Indels 1; Gaps 1;

Qy 1 CDLPQTHSLGSRRTLMLLAQMRRISLFSCLKDRHDFGFPQEEF...GNQFQKAETIPVLHE 59
Db 24 CDLPQTHSLGNRRALILLAQMGRISPFSCLKDRHDFGFPEEFD...GNQFQKAQAISVLHE 82

Qy 60 IQQIFNLFSTKDSSAAWDETLLDKFYIELYQQLNDLEACVIQGVGVTETPLMKEDSILAV 119
Db 83 IQQTFNLFSTEDSSAAWEQSLLEKFSIELYQQLNDLEACVIQEVGVEETPLMNVDSILAV 142

Qy 120 RKYFQRITLYLTEKKYSPCAWEVVRAEIMRSFSLSTNLQERLRRKD 165
Db 143 RKYFQRITLYLTEKKYSPCAWEVVRAEIMRSLSFSTNLQKRLRRKD 189

RESULT 11
I56313
interferon alpha 21 - human
C;Species: Homo sapiens (man)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 14-Jul-1999
C;Accession: I56313
R;Gren, E.; Berzin, V.M.; Tanenno, I.; Tsimanis, A.; Vishnevsky, V.; Apsalons, U.
J. Interferon Res. 4, 609-617, 1984
A;Title: Novel human leukocyte interferon subtype and structural comparison of eight int...
A;Reference number: I56313; MUID:85056523; PMID:6548765
A;Accession: I56313
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-181 <RES>
A;Cross-references: GB:M28596; NID:g184636; PIDN:AAA52512.1; PID:g184637
C;Genetics:
A;Gene: GDB:IFNA21
A;Cross-references: GDB:136360; OMIM:147584
A;Map position: 9p22-9p22
C;Superfamily: interferon alpha

Query Match 81.4%; Score 692.5; DB 2; Length 181;
Best Local Similarity 83.1%; Pred. No. 8.5e-57;
Matches 138; Conservative 8; Mismatches 19; Indels 1; Gaps 1;

Qy 1 CDLPQTHSLGSRRTLMLLAQMRRISLFSCLKDRHDFGFPQEEF...GNQFQKAETIPVLHE 59
Db 16 CDLPQTHSLGNRRALILLAQMGRISPFSCLKDRHDFGFPEEFD...GNQFQKAQAISVLHE 75

Qy 60 IQQIFNLFSTKDSSAAWDETLLDKFYIELYQQLNDLEACVIQGVGVTETPLMKEDSILAV 119
Db 76 IQQIFNLFSTEDSSAAWEQSLLEKFSIELYQQLNDLEACVIQEVGVEETPLMNVDSILAV 135

Qy 120 RKYFQRITLYLTEKKYSPCAWEVVRAEIMRSFSLSTNLQERLRRKE 164
Db 136 KKYLQRITLYLEKKYSPCAWEVVRAEIMRSFSLSKNLQERLRRKE 180

RESULT 12
I84464
interferon alpha F - human
C;Species: Homo sapiens (man)
C;Date: 02-Apr-1999 #sequence_revision 02-Apr-1999 #text_change 09-Jul-1999
C;Accession: I84464; I37585
R;Gren, E.Y.; Berzin, V.M.; [illegible]; Apsalons, U.; [illegible]
[illegible]; Lacha, V.P.; Kozlova, V.M.; [illegible]
Dokl. Biochem. 269, 47-95, 1983
A;Title: A new type of leukocyte interferon
A;Reference number: I37585
A;Accession: I84464
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-189 <RES>
A;Cross-references: GB:X76941; NID:g1...; PIDN:CAA54...; PID:g...
A;Accession: I37585
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-189 <RES>
A;Cross-references: GB:X76941; NID:g...; PIDN:CAA54...; PID:g...
C;Genetics:
A;Gene: IFNA
C;Superfamily: interferon alpha

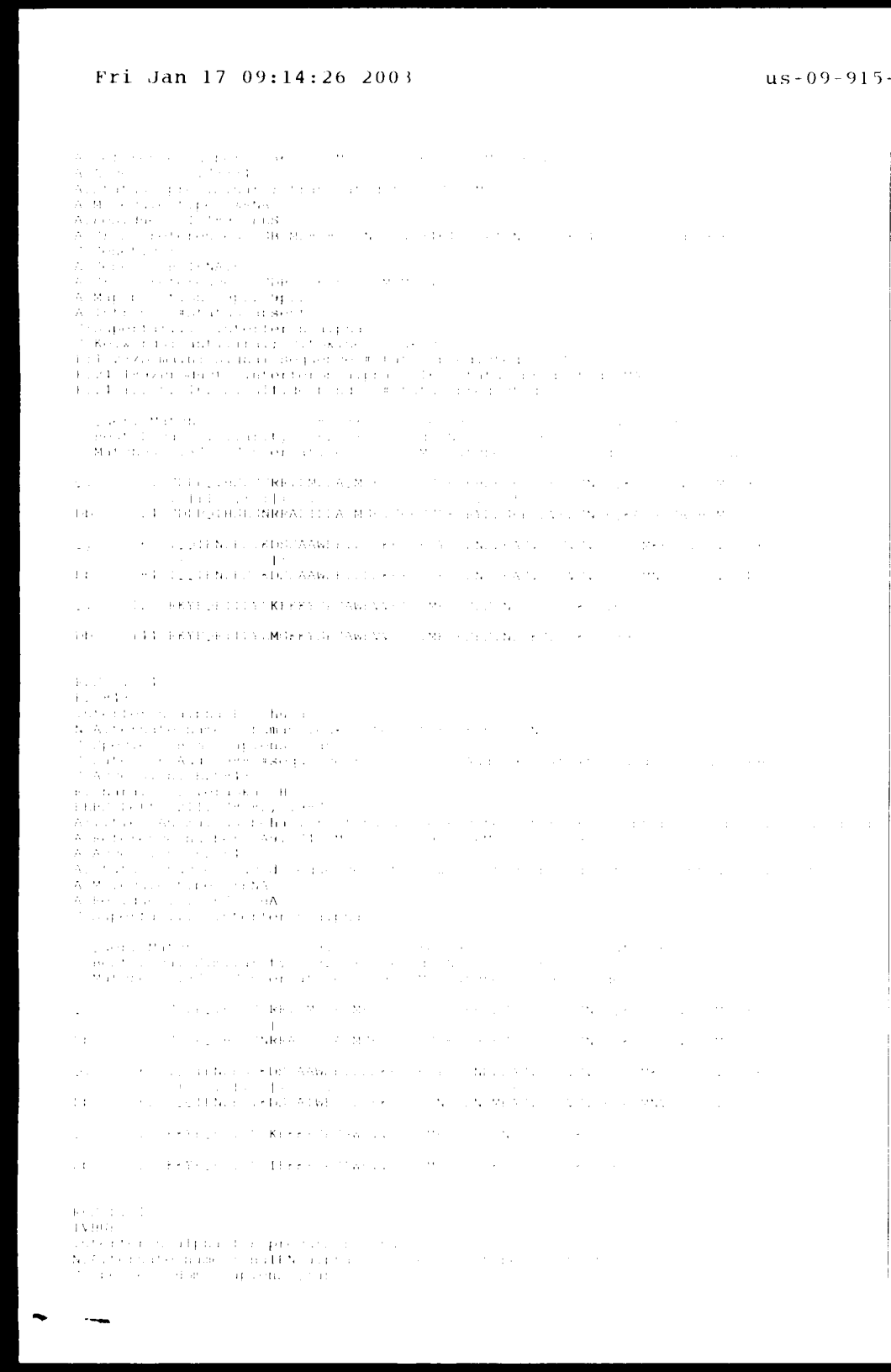
Query Match [illegible]; Score [illegible]; DB 2; Length [illegible];
Best Local Similarity [illegible]; Pred. No. [illegible];
Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

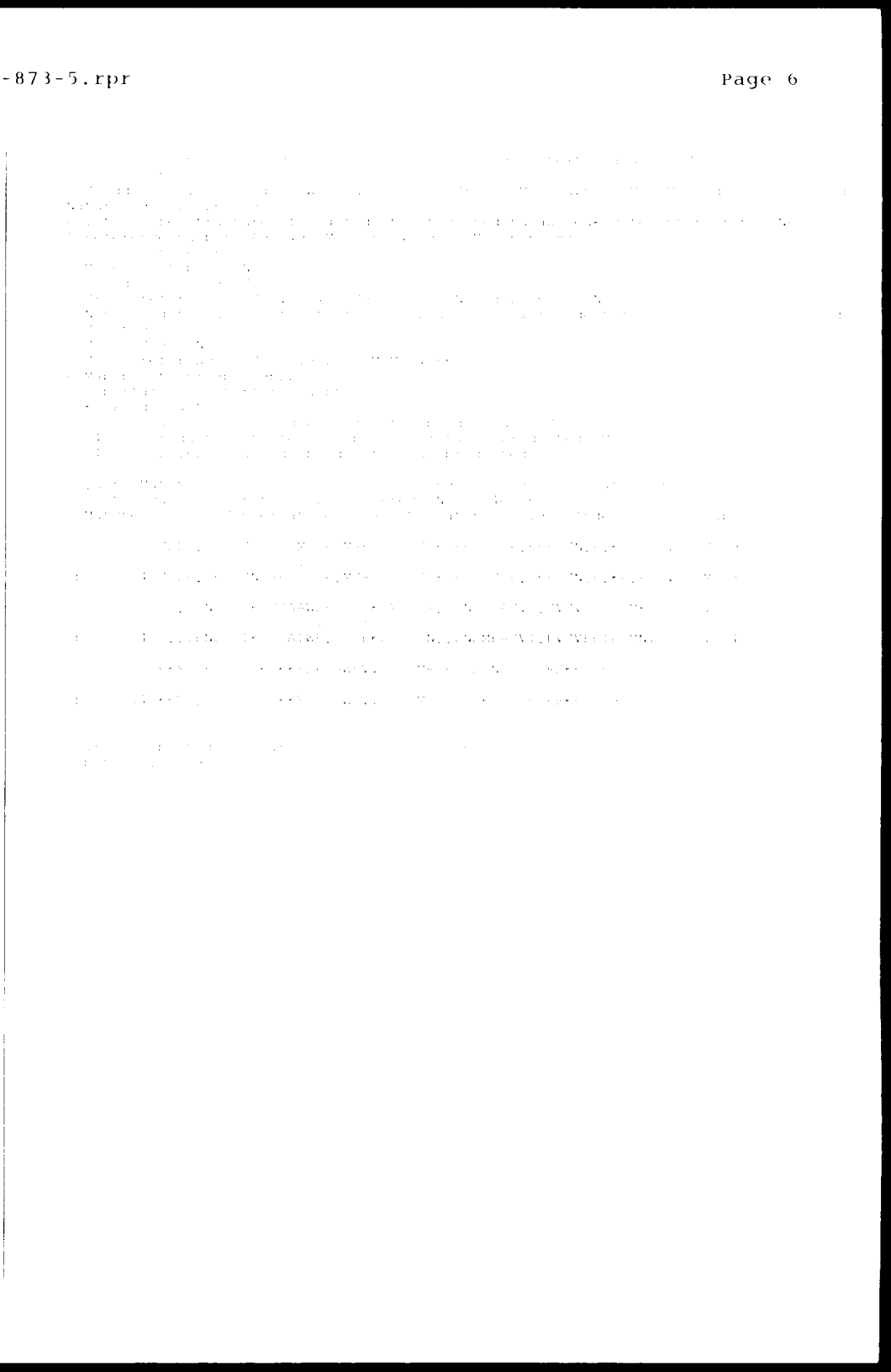
Qy 1 CDLPQTHSLGSRRTLMLLAQMRRISLFSCLKDRHDFGFPQEEF...GNQFQKAETIPVLHE 59
Db 24 CDLPQTHSLGNRRALILLAQMGRISPFSCLKDRHDFGFPQELD...GNQFQKAQAISVLHE 83

Qy 60 IQQIFNLFSTKDSSAAWDETLLDKFYIELYQQLNDLEACVIQGVGVTETPLMKEDSILAV 119
Db 84 IQQTFNLFSTEDSSAAWDQSLLEKFSIELYQQLNDLEACVIQEVGVEETPLMNVDSILAV 143

Qy 120 RKYFQRITLYLTEKKYSPCAWEVVRAEIMRSFSLSTNLQERLRRKE 164
Db 144 RKYFQRITLYLTEKKYSPCAWEVVRAEIMRSLSFSTNLQKRLRRKD 188

RESULT 13
[illegible]
interferon alpha-I-4b precursor - human
N;Alternate names: B-IFN-alpha I [illegible]; Interferon alpha I WA [illegible]
C;Species: Homo sapiens (man)
C;Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change [illegible]
C;Accession: S25454; AZ2048; I37585
R;[illegible]; Berzin, V.M.; [illegible]; Apsalons, A.A.; Apsite, [illegible]; Viksna, [illegible]
J. Mol. Biol. 185, 227-260, 1985
A;Title: Structural relationship of human interferon alpha genes and pseudogenes
A;Reference number: A-2916; MUID:86037205; PMID:4057246
A;Accession: S25454
A;Molecule type: DNA
A;Residues: 1-189 <HEN>
A;Cross-references: GB:X02955; NID:g32656; PIDN:CAA26731.1; PID:g32657
Krzyzynski, P.M.; Puke, B.; Berzins, A.;
Proc. Natl. Acad. Sci. U.S.A. 81, [illegible], 1984
A;Title: Human genomic library screened with IF [illegible]
[illegible]
A;Accession: AZ2048
A;Molecule type: DNA
A;Residues: 1-189 <KRZ>
A;Cross-references: GB:K...; NID:g...; PIDN:AAA...; PID:g...
F;Gren, E.; Berzin, V.M.; [illegible]; Apsalons, A.; Vishnevsky, V.; [illegible]
J. Interferon Res. 4, [illegible], 1984
A;Title: Novel human leukocyte interferon subtype and structural [illegible]





GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: January 17, 2003, 05:23:45 ; Search time 11.6456 Seconds
(without alignments)
585.646 Million cell updates/sec

Title: US-09-915-873-5
Perfect score: 851
Sequence: 1 CDLPQTHSLGSRRTLMLLAQ.....EIMRGTSLTINLQESLFSKE 165

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

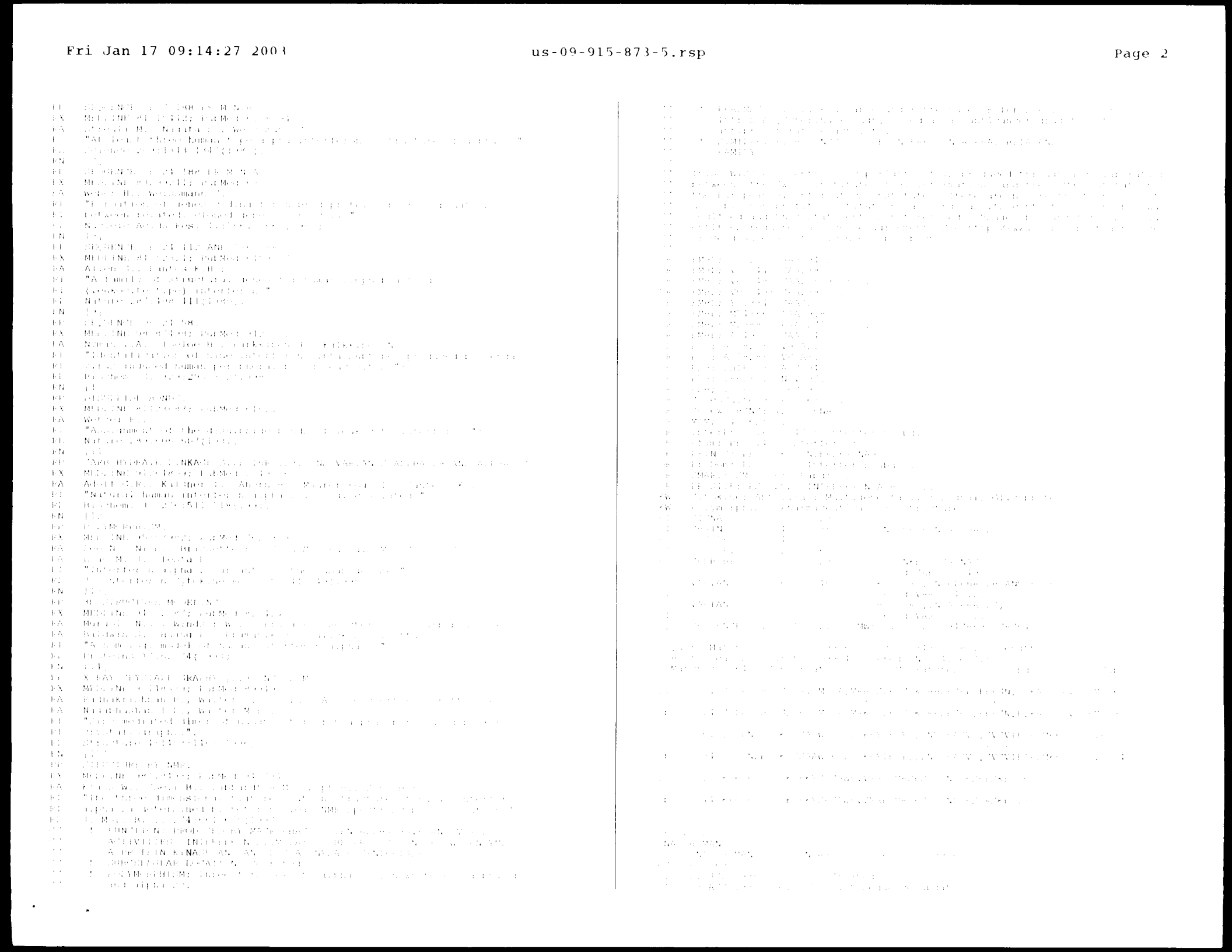
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	848	99.6	198	1	INA2_HUMAN	P01566 homo sapien
2	728.5	85.6	189	1	INA5_HUMAN	P01569 homo sapien
3	721.5	84.8	189	1	INA6_HUMAN	P05881 homo sapien
4	701.5	82.4	189	1	INAD_HUMAN	P01570 homo sapien
5	694.5	81.6	189	1	INA4_HUMAN	P05014 homo sapien
6	694.5	81.6	189	1	INAG_HUMAN	P01571 homo sapien
7	692.5	81.4	189	1	INAK_HUMAN	P01574 homo sapien
8	691.5	81.3	189	1	INAF_HUMAN	P05015 homo sapien
9	690.5	81.1	189	1	INAA_HUMAN	P01566 homo sapien
10	688.5	80.9	189	1	INA1_HUMAN	P01562 homo sapien
11	684.5	80.4	189	1	INA8_HUMAN	P32881 homo sapien
12	678.5	79.7	189	1	INA7_HUMAN	P01567 homo sapien
13	647.5	74.9	184	1	INA4_HORSE	P05006 equus cabal
14	633.5	74.4	184	1	INA2_HORSE	P05004 equus cabal
15	632.5	74.3	184	1	INA1_HORSE	P05003 equus cabal
16	632.5	74.3	184	1	INA3_HORSE	P05005 equus cabal
17	577.5	67.9	189	1	INA1_PIG	P49879 sus scrofa
18	562.5	66.1	189	1	INA5_MOUSE	P07319 mus musculu
19	547.5	64.3	189	1	INA1_MOUSE	P01572 mus musculu
20	534.5	62.7	189	1	INAH_BOVIN	P49828 bos taurus
21	531.5	62.5	189	1	INAA_BOVIN	P05007 bos taurus
22	530.5	62.3	189	1	INA1_BOVIN	P07148 bos taurus
23	528.5	62.1	189	1	INAB_BOVIN	P05008 bos taurus
24	525.5	61.8	189	1	INAC_BOVIN	P05009 bos taurus
25	525.5	61.8	189	1	INAD_BOVIN	P05010 bos taurus
26	525.5	61.8	190	1	INA7_MOUSE	P06799 mus musculu
27	525.5	61.8	192	1	INA1_RAT	P05011 rattus norv
28	524.5	61.6	190	1	INA2_MOUSE	P01573 mus musculu
29	523.5	61.5	190	1	INA9_MOUSE	P39235 mus musculu
30	516.5	60.7	189	1	INAG_BOVIN	P49877 bos taurus
31	515.5	60.6	189	1	INAF_BOVIN	P49876 bos taurus
32	511.5	60.1	189	1	INA6_MOUSE	P07450 mus musculu
33	509.5	59.9	189	1	INA8_MOUSE	P17660 mus musculu
34	508.5	59.7	189	1	INA_CHICK	P01575 gallus gall
35	488	57.3	189	1	INA1_MACFA	P48616 macaca fasc
36	484.5	56.9	189	1	INA_CHICK	P01576 gallus gall
37	454	53.3	198	1	INA_CHICK	P01577 gallus gall
38	446.5	52.5	189	1	INA_BOVIN	P01568 bos taurus
39	437.5	51.4	172	1	INA_MOUSE	P07320 mus musculu
40	434.5	51.1	198	1	INA_HUMAN	P01565 homo sapien
41	434.5	51.1	149	1	INA_BOVIN	P01567 bos taurus
42	433.5	51.0	172	1	INA_MOUSE	P07321 mus musculu
43	429.5	50.5	165	1	INA_RAT	P01573 rattus norv
44	426.5	50.1	167	1	INA_RABIT	P01574 oryctolagus
45	426.5	50.1	167	1	INA_RABIT	P01575 oryctolagus

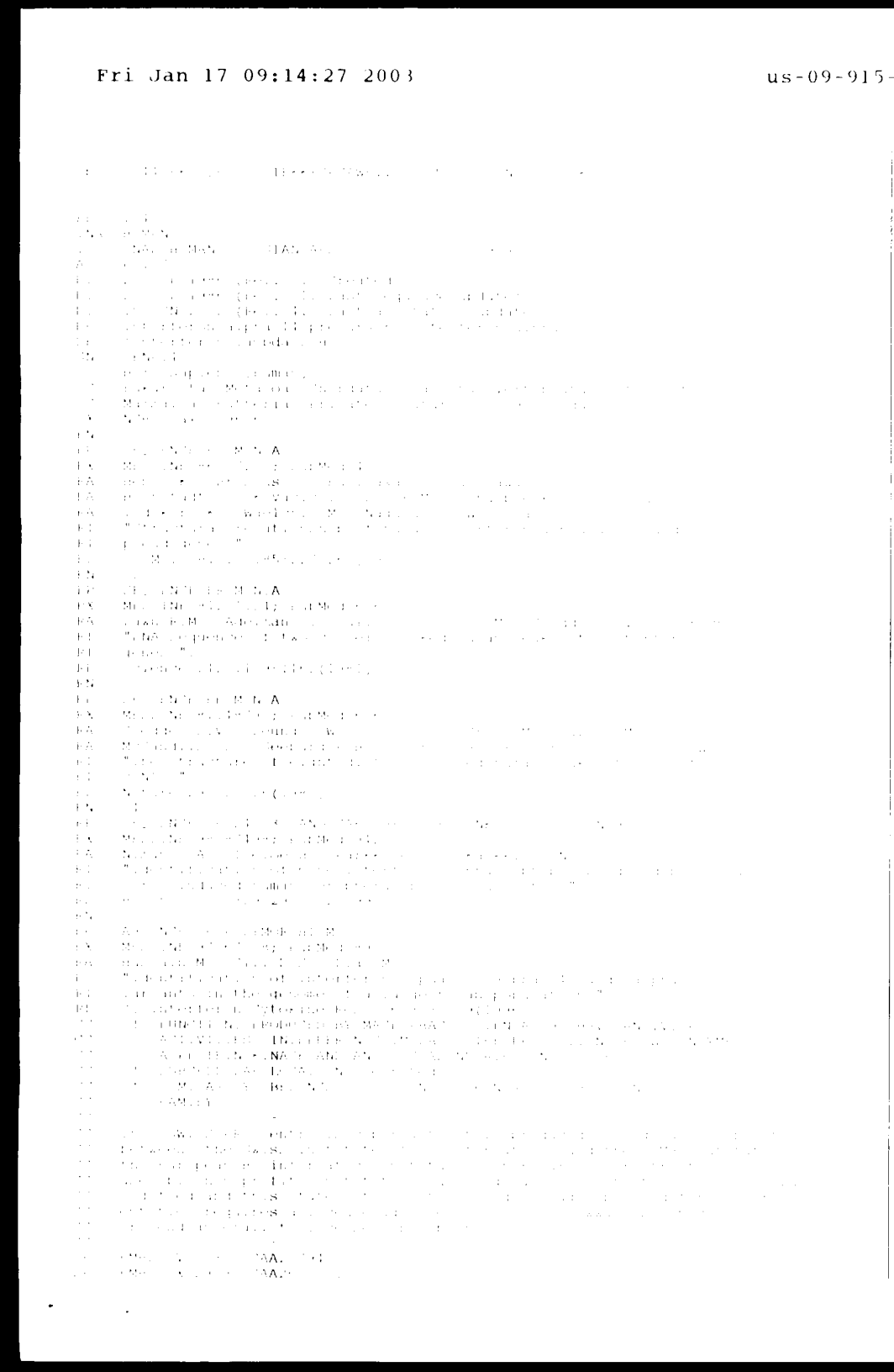
ALIGNMENTS

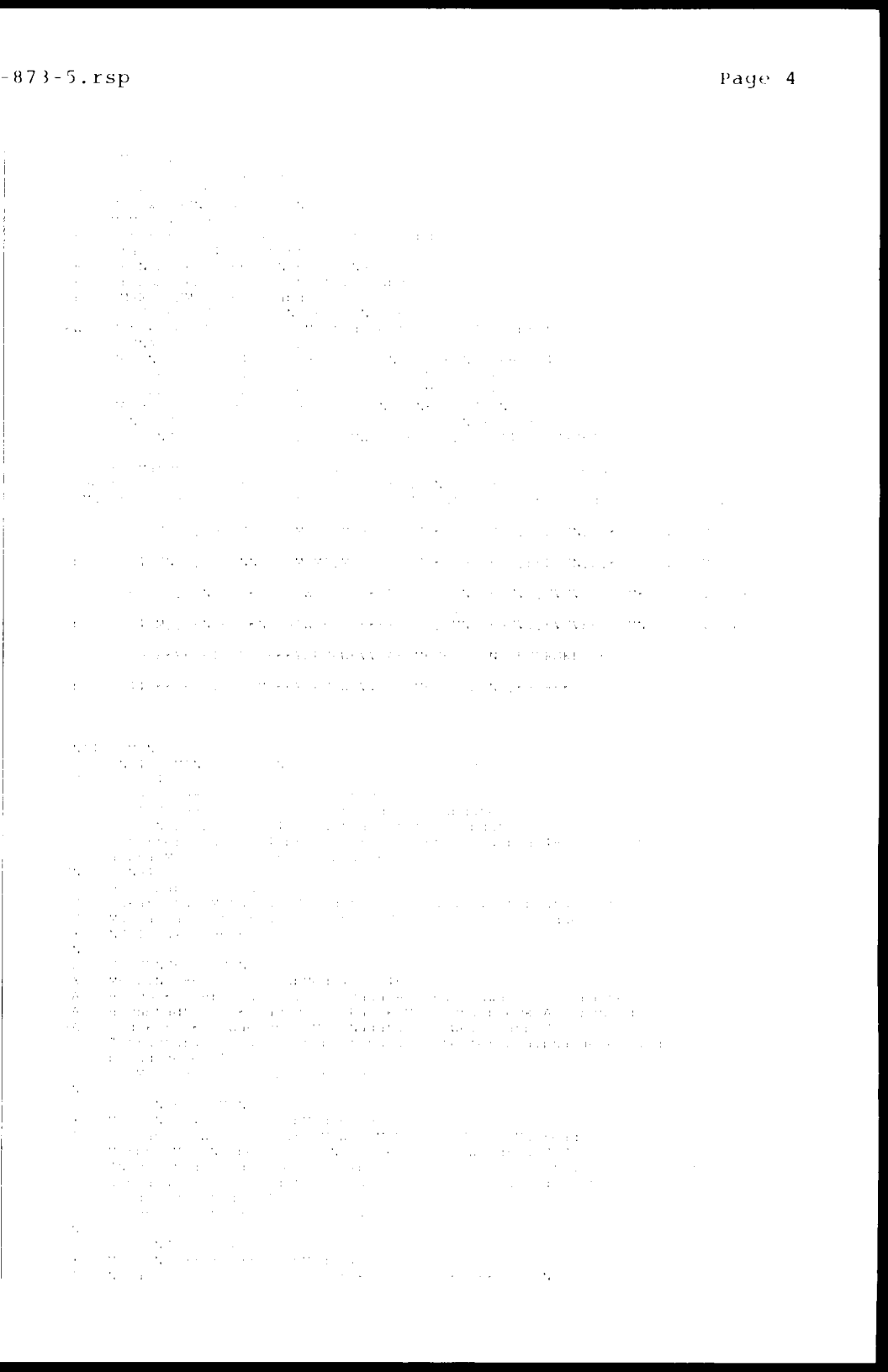
RESULT 1
INA2_HUMAN
ID INA2_HUMAN STANDARD; PRT; 198 AA.
AC P01566; P01564; Q14099;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JAN-2002 (Rel. 41, Last annotation update)
DE Interferon alpha A precursor (Interferon alpha A) (IFN-A).
GN IFNA2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=81052222; PubMed=6165748;
RA Goeddel D.V., Yelverton E., Ullrich A., Heyneker H.L., Miozzari G.,
RA Holmes W., Seeburg P.H., Hobom G., Taira T., Ramachandran R.,
RA Mayda S., McCandliss R., Sloan A., Lai C.-Y., Stebbing N.,
RA Familletti P.C., Pestka S.;
RT "Human leukocyte interferon produced in E. coli is biologically
RT active.";
RL Nature 287:411-416(1980).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=81148795; PubMed=6787156;
RA Goeddel D.V., Leung D.W., Dull T.J., Gross M., Lawn R.M.,
RA McCandliss R., Seeburg P.H., Ullrich A., Yelverton E., Gray P.W.;
RT "The structure of eight distinct cloned human leukocyte interferon
RT cDNAs.";
RL Nature 290:20-26(1981).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=82061261; PubMed=6172331;
RA Lawn R.M., Gross M., Houck C.M., Franke A.E., Gray P.W.,
RA Goeddel D.V.;
RT "DNA sequence of a major human leukocyte interferon gene.";
RL Proc. Natl. Acad. Sci. U.S.A. 78:5435-5439(1981).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Bone marrow tumor;
RX MEDLINE=86089761; PubMed=3000855;
RA Oliver G., Ballas P., Valle F., Gideon X., Beaujean D.,
RT "Cloning of human leukocyte interferon cDNA and a strategy for
RT production in E. coli.";
RL Res. Inflamm. Microbiol. 2:111-120(1985).
RN [5]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=98357449; PubMed=9687510;
RA Arnheiter E., Bandle C., Carbonell N., Mar [illegible]., Brial E., Lata [illegible].
RA Mantoni F., Chabannon C.;
RT "A defective retroviral [illegible] transduce human interferon [illegible]
RT transduce human leukemic cell lines.";
RL Cancer Gene Ther. 5:247-252(1998).
RN [6]



```
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Interferon alpha-5 precursor (Interferon alpha-C) (Leif C) (Interferon
DE   alpha-61).
GN   IFNA5.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86037205; PubMed=4057246;
RA   Henco K., Brosius J., Fujisawa A., Fujisawa J.-I., Haynes J.R.,
RA   Hochstadt J., Kovacic T., Pasek M., Schambock A., Schmid J.,
RA   Tadokoro K., Waelchli M., Nagata S., Weissmann C.;
RT   "Structural relationship of human interferon alpha genes and
RT   pseudogenes.";
RL   J. Mol. Biol. 185:227-260(1985).
RN   [2]
RP   SEQUENCE OF 57-189 FROM N.A.
RC   TISSUE=Spleen;
RX   MEDLINE=81148795; PubMed=6163083;
RA   Goeddel D.V., Leung D.W., Dull T.J., Gross M., Lawn R.M.,
RA   McCandliss R., Seeburg P.H., Ullrich A., Yelverton E., Gray P.W.;
RT   "The structure of eight distinct cloned human leukocyte interferon
RT   cDNAs.";
RL   Nature 290:20-26(1981).
CC   -!- FUNCTION: PRODUCED BY MACROPHAGES, IFN-ALPHA HAVE ANTIVIRAL
CC       ACTIVITIES. INTERFERON STIMULATES THE PRODUCTION OF
CC       A PROTEIN KINASE AND AN OLIGOADENYLATE SYNTHETASE.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: BELONGS TO THE INTERFERON ALPHA, BETA AND DELTA
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X02956; CAA26702.1; -.
DR   EMBL; V00541; CAA22802.1; -.
DR   PIR; A01833; IVHUA7.
DR   HSSP; P01563; 2HIE.
DR   Genew; HGNC:5426; IFNA5.
DR   MIM; 147565; -.
DR   InterPro; IPR000471; Interferon_abd.
DR   Pfam; PF00143; interferon; 1.
DR   PRINTS; PR00266; INTERFERONAB.
DR   ProDom; PD000550; Interferon_abd; 1.
DR   SMART; SM00076; IFabd; 1.
DR   PROSITE; PS00252; INTERFERON_A_B_D; 1.
KW   Cytokine; Antiviral; Multigene family; Signal.
FT   SIGNAL        1     23
FT   CHAIN        24    189       INTERFERON ALPHA 5.
FT   DISULFID     24    122       BY SIMILARITY.
FT   DISULFID     52    162       BY SIMILARITY.
SQ   SEQUENCE   189 AA;  21942 MW;  [illegible] CRC64;

  Query Match             85.6%;  Score 728.5;  DB 1;  Length 189;
  Best Local Similarity   85.5%;  Pred. No. 4.4e-60;
  Matches  142;  Conservative    7;  Mismatches   16;  Indels    1;  Gaps    1;

Qy       1 CDLPQTHSLGSRRTLMLLAQMRRISLFSCLKDRHDFGFPQEEF DNKQFQKAPAISVLHELI 59
           ||||||||| :||||||:||| ||| |||||||||||||||||| |||||||: | |||||
Db      24 CDLPQTHSLSNRRTLMMAQMGRISPFSCLKDRHDFRFPQEEF DGNQFQKAQAISVLHEMI 83

Qy      60 QQTFNLFSTKDSSAAWDETLLDKFYTELYQQLNDLEACVIQGVGVTETPLMKEDSILAV 119
           ||| |||||||||| |||||||||||||||||||||||:| ||| :|||| |||| |
Db      84 QQTFNLFSTKDSSATWDETLLDKFYTELYQQLNDLEACVMQEVGVIETPLMNVDSILAV 143

Qy     120 RKYFQRITLYLTEKKYSPCAWEVVRAEIMRSFSLSTNLQ 157
```

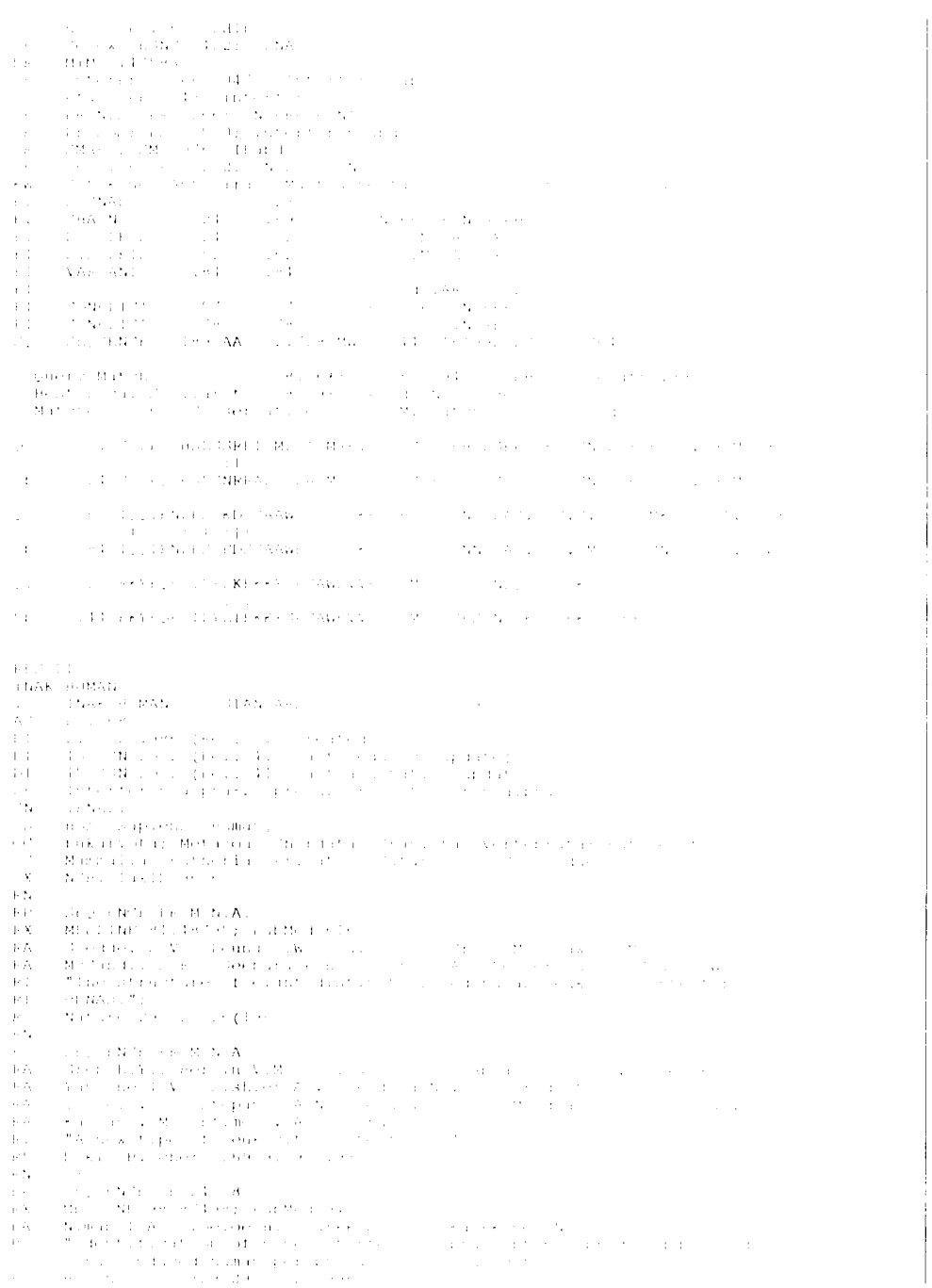
```
          ||||||||||| |||| |||||||| |||||| |||| |||| ||
Db     144 RKYFQRITLYLTEKKYSPCAWEVVRAEIMRSFSLSTNLQ 181


RESULT 3
IFNA5_HUMAN
ID   IFNA5_HUMAN     STANDARD;      PRT;   189 AA.
AC   P05013;
DT   13-AUG-1987 (Rel. 05, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Interferon alpha-5 precursor (Interferon alpha-C) (Leif C) (Interferon
DE   alpha-54).
GN   IFNA5.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86037205; PubMed=4057246;
RA   Henco K., Brosius J., Fujisawa A., Fujisawa J.-I., Haynes J.R.,
RA   Hochstadt J., Kovacic T., Pasek M., Schambock A., Schmid J.,
RA   Tadokoro K., Waelchli M., Nagata S., Weissmann C.;
RT   "Structural relationship of human interferon alpha genes and
RT   pseudogenes.";
RL   J. Mol. Biol. 185:227-260(1985).
RN   [2]
RP   SEQUENCE OF 57-189 FROM N.A.
RC   TISSUE=Spleen;
RX   MEDLINE=81148795; PubMed=6163083;
RA   Goeddel D.V., Leung D.W., Dull T.J., Gross M., Lawn R.M.,
RA   McCandliss R., Seeburg P.H., Ullrich A., Yelverton E., Gray P.W.;
RT   "The structure of eight distinct cloned human leukocyte interferon
RT   cDNAs.";
RL   Nature 290:20-26(1981).
CC   -!- FUNCTION: PRODUCED BY MACROPHAGES, IFN-ALPHA HAVE ANTIVIRAL
CC       ACTIVITIES. INTERFERON STIMULATES THE PRODUCTION OF TWO ENZYMES:
CC       A PROTEIN KINASE AND AN OLIGOADENYLATE SYNTHETASE.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: BELONGS TO THE INTERFERON ALPHA, BETA AND DELTA
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X02956; CAA26704.1; -.
DR   PIR; A23754; IVHUF4.
DR   HSSP; P01563; 2HIE.
DR   Genew; HGNC:5427; IFNA6.
DR   MIM; 147566; -.
DR   InterPro; IPR000471; Interferon_abd.
DR   Pfam; PF00143; interferon; 1.
DR   PRINTS; PR00266; INTERFERONAB.
DR   ProDom; PD000550; Interferon_abd; 1.
DR   SMART; SM00076; IFabd; 1.
DR   PROSITE; PS00252; INTERFERON_A_B_D; 1.
KW   Cytokine; Antiviral; Multigene family; Signal.
FT   SIGNAL        1     23
FT   CHAIN        24    189       INTERFERON ALPHA-54.
FT   DISULFID     24    122       BY SIMILARITY.
FT   DISULFID     52    162       BY SIMILARITY.
SQ   SEQUENCE   189 AA;  [illegible] MW;  [illegible] CRC64;

  Query Match             84.6%;  Score [illegible];  DB 1;  Length [illegible];
  Best Local Similarity   86.7%;  Pred. No. [illegible];
  Matches  144;  Conservative    4;  Mismatches   18;  Indels    1;  Gaps    1;

Qy       1 CDLPQTHSLGSRRTLMLLAQMRRISLFSCLKDRHDFGFPQEEF DNKQFQKAPAISVLHELI 59
           ||||||||| |||:||||||||||||||||||||||||||||| |||||| |||| | |||||
Db      24 CDLPETHSLRNRRALMILAQMSRIAPSSFLKDLNKRFTFGLLE DSRLQTNEQAASVMHNLH 83

Qy      60 QQTFNLFSTKDSSAAWDETLLDKFYTELYQQLNDLEACVIQGVGVTETPLMKEDSILAV 119
           ||| |||||||||| |||||||||||||||||||||||:| ||| :|||| |||| |
Db      84 QQIFSLFHTERSSAAWNMTLLDQLCTGLHQQINHLQTCLKPQVGLEVTEYEDQSLLGV 143

Qy     120 RKYFQRITLYLTEKKYSPCAWEVVRAEIMRSFSLSTNLQ 157
           ||||||||||| |||| |||||||| |||||| |||| ||
Db     144 RKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSKNFQ 181
```





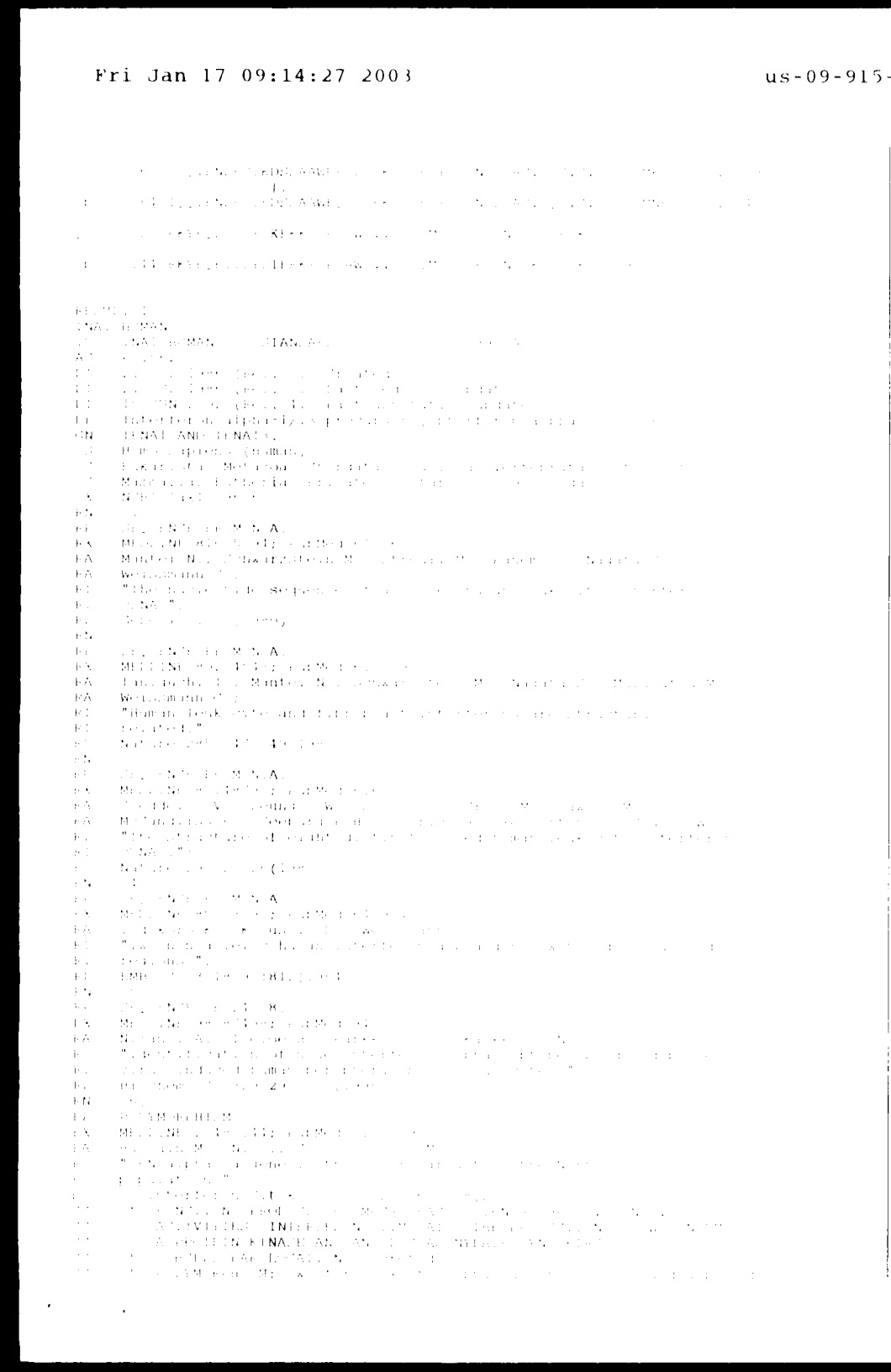
```
RT   "Identification of nine interferon-alpha subtypes produced by Sendai
RT   virus-induced human peripheral blood leucocytes.";
RL   Biochem. J. 329:295-302(1998).
RN   [4]
RP   POLYMORPHISM.
RX   MEDLINE=97474410; PubMed=9425434;
RA   Hussain M., Gill D.S., Liao M.-J.;
RT   "Both variant forms of interferon-alpha4 gene (IFNA4a and IFNA4b) are
RT   present in the human population.";
RL   J. Interferon Cytokine Res. 17:559-566(1997).
CC   -!- FUNCTION: PRODUCED BY MACROPHAGES, IFN-ALPHA HAVE ANTIVIRAL
CC       ACTIVITIES. INTERFERON STIMULATES THE PRODUCTION OF TWO ENZYMES:
CC       A PROTEIN KINASE AND AN OLIGOADENYLATE SYNTHETASE.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- POLYMORPHISM: Two forms exist: alpha-4a and alpha-4b (shown here).
CC       They seem to be equally abundant.
CC   -!- SIMILARITY: BELONGS TO THE INTERFERON ALPHA, BETA AND DELTA
CC       FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaborative
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X02955; CAA26701.1; -.
DR   EMBL; M27318; AAA52726.1; -.
DR   PIR; E23753; IVHU4b.
DR   HSSP; P01563; 2HIE.
DR   Genew; HGNC:5425; IFNA4.
DR   MIM; 147564; -.
DR   InterPro; IPR000471; Interferon_alpd.
DR   Pfam; PF00143; interferon; 1.
DR   PRINTS; PR00266; INTERFERONAB.
DR   ProDom; PD000550; Interferon_alpd; 1.
DR   SMART; SM00076; IFabd; 1.
DR   PROSITE; PS00252; INTERFERON_A_B_D; 1.
KW   Cytokine; Antiviral; Multigene family; Polymorphism; Signal.
FT   SIGNAL       1     23
FT   CHAIN       24    189       INTERFERON ALPHA 4.
FT   DISULFID    24    122       BY SIMILARITY.
FT   DISULFID    52    162       BY SIMILARITY.
FT   VARIANT     74     74       I -> A (IN ALPHA-4A).
FT                               /FTId=VAR_013002.
FT   VARIANT    137    137       V -> E (IN ALPHA-4A).
FT                               /FTId=VAR_013003.
SQ   SEQUENCE   189 AA;  21903 MW;  4198EF9CBEEAE8D8 CRC64;

  Query Match             81.6%;  Score 694.5;  DB 1;  Length 189;
  Best Local Similarity   81.9%;  Pred. No. 5.8e-57;
  Matches  136;  Conservative   12;  Mismatches   17;  Indels    1;  Gaps    1;

Qy       1 CDLPQTHSLGSRRTLMLLAQMRRISPFSCLKDRHDFGFPQEEFDGNQFQKAPAISVLHEL 60
           ||||||||||:|| |:|||| ||| |||||||||||||||| |:|||| : |:||||
Db      24 CDLPQTHSLGNRRALILLAQMGRISHFSCLKDRHDFEFPEEEFDGHQFQKTQAISVLHEM 83

Qy      61 IQQIFNLFSTKDSSAAWDETLLDKFYIELYQQLNDLEACVIQGVGVTETPLMKEDSILAV 120
           ||| |||||:|||||||:::|||:|| |||||||||||||||| ||| |||||  ||||||
Db      84 IQQTFNIFSTFDSSAAWFQSLLEKFSTELYQQLNDLEACVIQEVGVEETPLMNEDSILAV 143

Qy     121 RKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTNLQESLRSKE 165
           |||||||||| |||||||||||||||| | ||||: || ||
Db     144 RKYFQRITVYLIEKKYSPCAWPVVRAFIMRSLSFSTNLQKRLRRK- 188


RESULT 6
INAG_HUMAN
ID   INAG_HUMAN     STANDARD;      PRT;   189 AA.
AC   P01571; Q14639;
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
```

```
DT   [5] N-terminal [...] (Rel. [...], [...] of [...] [...])
DE   Interferon alpha-4 precursor (IFN-alpha 4) (IFN-
DE   alpha4) (Interferon alpha-60).
GN   IFNA4.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=81201121; PubMed=6163502;
RA   Lawn R.M., Adelman J., Bell S.R., Collen D., Goeddel D.V., Grant P.M., Gross M., Haber J.E., Hoyer B.H., Maniatis T., Nagata S., Nathans D., Nishi S., Ohno S., Pestka S., Ruddle F.H., Shapiro L., Shaw G., Tabor H., Taylor J.M., Weissmann C., Yelverton E.;
RT   "DNA sequences of two closely linked human leukocyte interferon
RT   genes.";
RL   Science 212:1159-1162(1981).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=82227[illegible]; PubMed=[illegible];
RA   Mizoguchi J., Titka P.M., Ito N.K.;
RT   "Efficient expression in Escherichia coli of two species [illegible]
RT   interferon-alpha and their [illegible]";
RL   DNA 1:[illegible]-[illegible].
RN   [3]
RP   SEQUENCE OF 11-189 FROM N.A.
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   Lundblad R.L., Abildgaard F., Jacobsen H., Kakkar V.V.;
RT   "Differential expression of interferon alpha genes in a subline of
RT   Namalva cells.";
RL   J. Interferon Res. [illegible]-[illegible].
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87264[illegible]; PubMed=[illegible];
RA   Sarvetnick N., Shichiri M., [illegible], Watanabe Y.,
RA   Yelton D.E., Dembrow D.N., Fox H.S., Notkins A.L.;
RA   Gazaryan K.G., Monastyrskaya E.S., Scarlato E.L.;
RT   "Cloning and the determination of the nucleotide sequence of a human
RT   [illegible] human leukocyte interferon gene].";
RL   Antibiot. Med. Biotekhnol. [illegible]-[illegible](1986).
RN   [5]
RP   SEQUENCE OF 24-58.
RX   MEDLINE=98268[illegible]; PubMed=[illegible];
RA   Nyman T.A., Lehtoo E., Parkkinen J., Kalkkinen N.;
RT   "Identification of nine interferon alpha subtypes produced by Sendai
RT   virus induced human peripheral blood leucocytes.";
RL   Biochem. J. [illegible]-[illegible](1998).
RN   [6]
RP   VARIANT ARG-[illegible].
RX   MEDLINE=98376[illegible]; PubMed=[illegible];
RA   Hussain M., Gill D.S., [illegible], Liao M.-J.;
RT   "A new allele of interferon alpha4 gene (IFNA4) differ IFN activity
RT   major variant in human population.";
RL   J. Interferon Cytokine Res. 18:[illegible]-[illegible](1998).
CC   -!- FUNCTION: PRODUCED BY MACROPHAGES, IFN-ALPHA HAVE ANTIVIRAL
CC       ACTIVITIES. INTERFERON STIMULATES THE PRODUCTION OF TWO ENZYMES:
CC       A PROTEIN KINASE AND AN OLIGOADENYLATE SYNTHETASE.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: BELONGS TO THE INTERFERON ALPHA, BETA AND
CC       FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaborative
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M11[illegible]; AAA52720.1; -.
DR   EMBL; V00532; CAA23[illegible]; -.
DR   EMBL; M[illegible]; AAA52[illegible]; -.
DR   EMBL; M[illegible]; AAA52[illegible]; -.
DR   PIR; A01865; IVHUA5.
DR   PIR; A[illegible]; A[illegible].
```

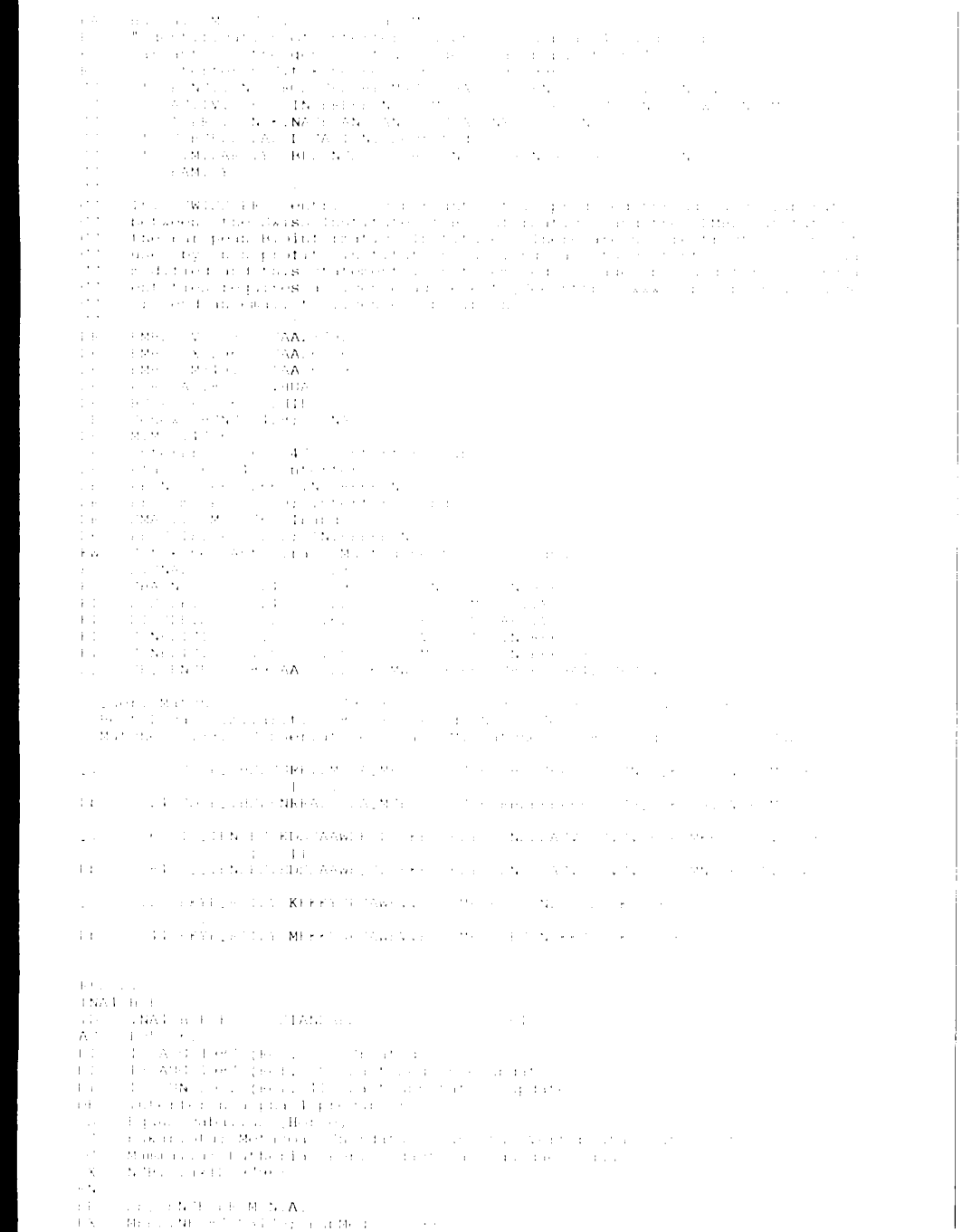


```
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86037205; PubMed=4057246;
RA   Henco K., Brosius J., Fujisawa A., Fujisawa J.-I., Haynes J.R.,
RA   Hochstadt J., Kovacic T., Pasek M., Schambock A., Schmid C.,
RA   Todokoro K., Waelchli M., Nagata S., Weissmann C.;
RT   "Structural relationship of human interferon alpha genes and
RT   pseudogenes.";
RL   J. Mol. Biol. 185:227-260(1985).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85038533; PubMed=6387705;
RA   Torczynski R.M., Fuke M., Bollon A.P.;
RT   "Human genomic library screened with 17-base oligonucleotide probes
RT   yields a novel interferon gene.";
RL   Proc. Natl. Acad. Sci. U.S.A. 81:6451-6455(1984).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85056523; PubMed=6548765;
RA   Gren E., Berzin V.M., Jansone I., Isimanis A., Vishnevsky Y.,
RA   Apsalons U.;
RT   "Novel human leukocyte interferon subtype and structural comparison
RT   of alpha interferon genes.";
RL   J. Interferon Res. 4:609-617(1984).
CC   -!- FUNCTION: PRODUCED BY MACROPHAGES. IFN ALPHA HAVE ANTIVIRAL
CC       ACTIVITIES. INTERFERON STIMULATES THE PRODUCTION OF TWO ENZYMES:
CC       A PROTEIN KINASE AND AN OLIGOADENYLATE SYNTHETASE.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: BELONGS TO THE INTERFERON ALPHA, BETA AND OMEGA
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch)
CC   ---------------------------------------------------------------------
DR   EMBL; X02957; CAA26703.1; -.
DR   EMBL; K02055; AAA52727.1; -.
DR   EMBL; M28585; AAA36042.1; -.
DR   PIR; G23753; IVHU16.
DR   PIR; A22068; A22068.
DR   HSSP; P01563; 2HIE.
DR   Genew; HGNC:5421; IFNA16.
DR   MIM; 147580; -.
DR   InterPro; IPR000471; Interferon_abd.
DR   Pfam; PF00143; interferon; 1.
DR   PRINTS; PR00266; INTERFERONA.
DR   ProDom; PD000550; Interferon_abd; 1.
DR   SMART; SM00076; IFabd; 1.
DR   PROSITE; PS00252; INTERFERON_A_B_D; 1.
KW   Cytokine; Antiviral; Multigene family; Signal.
FT   SIGNAL        1     23
FT   CHAIN        24    189       INTERFERON ALPHA-16.
FT   DISULFID     24    122       BY SIMILARITY.
FT   DISULFID     52    162       BY SIMILARITY.
SQ   SEQUENCE   189 AA;  21711 MW;  F0882EF787E1585F CRC64;

  Query Match             81.4%;  Score 691.5;  DB 1;  Length 189;
  Best Local Similarity   82.5%;  Pred. No. 1.1e-56;
  Matches  137;  Conservative    7;  Mismatches   21;  Indels    1;  Gaps    1;

Qy       1 CDLPQTHSLGNRRALILLAQMGRISPFSCLKDRHDFGFPQEVFDGNQFQKAQAISAFHEMM 60
           ||||||||||:|| |:|||||  ||| |||||||:|||||||  | |||||||: |   |||
Db      24 CDLPQTHSLGNRRALILLAQMGRISHFSCLKDRYDFGFPQEVFDGNQFQKAEAISAFHEMM 83

Qy      61 PQIFNLFSTKDSSAAWDETLLDKFYTELYQQLNDLEACVIQGVGVTETPLMKEDSILAV 119
           ||| |||||||||||||||||||| ||:|||||||:|| |||||||||||||:|||||||
Db      84 PQQTFNLFSTKDSSAAWDETLLDKFYIELFQQLNDLEACVIQEVGVEETPLMNEDSILAV 143

Qy     120 RKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTNLQESLRSKE 165
```

```
RESULT 2
INAA_HUMAN
ID   INAA_HUMAN     STANDARD;      PRT;   166 AA.
AC   P01566;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Interferon alpha-10 precursor (Interferon alpha") (IFN-").
DE   (Interferon alpha-6L).
GN   IFNA10.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=84117470; PubMed=6547765;
RA   Goeddel D.V., Leung D.W., Dull T.J., Gross M., Lawn R.M.,
RA   McCandliss R., Seeburg P.H., Ullrich A., Yelverton E., Gray P.W.;
RT   "The structure of eight human leukocyte interferon genes.";
RL   Nature 290:20-26(1981).
RN   [2]
RP   SEQUENCE OF 24-53.
RX   MEDLINE=96367788; PubMed=8752210;
RA   Nyman T.A., Parmio M., Pirkkiaho S., Kalkkinen N.;
RT   "Identification of nine O-glycosylation sites in human
RT   virus-induced human peripheral blood leukocytes.";
RL   Biochem. J. 317:265-272(1996).
CC   -!- FUNCTION: PRODUCED BY MACROPHAGES. IFN ALPHA HAVE ANTIVIRAL
CC       ACTIVITIES. INTERFERON STIMULATES THE PRODUCTION OF TWO ENZYMES:
CC       A PROTEIN KINASE AND AN OLIGOADENYLATE SYNTHETASE.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: BELONGS TO THE INTERFERON ALPHA, BETA AND OMEGA
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; V00551; CAA23812.1; -.
DR   PIR; A01566; IVHU6.
DR   HSSP; P01563; 2HIE.
DR   Genew; HGNC:5426; IFNA.
DR   MIM; 147570; -.
DR   InterPro; IPR000471; Interferon_abd.
DR   Pfam; PF00143; interferon; 1.
DR   PRINTS; PR00266; INTERFERONA.
DR   ProDom; PD000550; Interferon_abd; 1.
DR   SMART; SM00076; IFabd; 1.
DR   PROSITE; PS00252; INTERFERON_A_B_D; 1.
KW   Cytokine; Antiviral; Multigene family; Signal.
FT   SIGNAL        1     23
FT   CHAIN        24    166       INTERFERON ALPHA-10.
FT   DISULFID     24    122       BY SIMILARITY.
FT   DISULFID     52    162       BY SIMILARITY.
SQ   SEQUENCE   166 AA;  19000 MW;  [illegible] CRC64;

  Query Match             [illegible]  Score [illegible]  DB 1;  Length 166;
  Best Local Similarity   [illegible]  Pred. No. [illegible];
  Matches  [illegible]  Conservative  [illegible]  Mismatches  [illegible]  Indels  [illegible]  Gaps  [illegible];

Qy       1 CDLPQTHSLGNRRALILLAQMGRISPFSCLKDRHDFGFPQEVFDGNQFQKAQAISAFHEMM 60
           [illegible]
Db      21 CDLPQTHSLGNRRALILLAQMGRISPFSCLKDRHDFGFPQEVFDGNQFQKAQAISAFHEMM 80

Qy      61 PQIFNLFSTKDSSAAWDETLLDKFYTELYQQLNDLEACVIQGVGVTETPLMKEDSILAV 119
           [illegible]
Db      81 PQQFNLFSTKDSSAAWDETLLDKFYIELFQQMNDLEACVIQEVGVEETPLMNEDSILAV 140

Qy     120 RKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTNLQESLRSKE 165
```



```
RA   Henco K., Brosius J., Fujisawa A., Fujisawa J.-I., Haynes J.R.,
RA   Hochstadt T., Kawade T., Klett M., Kourilsky P., Maheux R.,
RA   Todokoro K., Waelchli M., Nagata S., Weissmann C.;
RT   "Structural relationship of human interferon alpha genes and
RT   pseudogenes.";
RL   J. Mol. Biol. 185:227-260(1985).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE 84193614, PubMed 6325333;
RA   Bowden D.W., Mao J., Gill T., Hsiao K., Lillquist J.S., Testa D.,
RA   Vovis G.F.;
RT   "Cloning of eukaryotic genes in single-strand phage vectors: the
RT   human interferon genes.";
RL   Gene 27:87-99(1984).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=81174739; PubMed 6164049;
RA   Yelverton E., Leung D., Week P., Gray P.W., Goeddel D.V.;
RT   "Bacterial synthesis of a novel human leukocyte interferon.";
RL   Nucleic Acids Res. 9:731-741(1981).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=81148795; PubMed 6163083;
RA   Goeddel D.V., Leung D.W., Dull T.J., Gross M., Lawn R.M.,
RA   McCandliss R., Seeburg P.H., Ullrich A., Yelverton E., Gray P.W.;
RT   "The structure of eight distinct cloned human leukocyte interferon
RT   cDNAs.";
RL   Nature 290:20-26(1981).
RN   [5]
RP   SEQUENCE OF 24-53.
RX   MEDLINE=98087498; PubMed 9425112;
RA   Nyman T.A., Doelow H., Parkkinen J., Kalkkinen N.;
RT   "Identification of nine interferon-alpha subtypes produced by Sendai
RT   virus-induced human peripheral blood leucocytes.";
RL   Biochem. J. 329:295-302(1998).
RN   [6]
RP   ABSENCE OF POLYMORPHISM.
RX   MEDLINE=96433939; PubMed=8836918;
RA   Hussain M., Gill D.S., Liao M.-J., Testa D.;
RT   "Interferon-alpha 8b is the only variant of interferon-alpha 8
RT   identified in a large human population.";
RL   J. Interferon Cytokine Res. 16:523-529(1996).
CC   -!- FUNCTION: PRODUCED BY MACROPHAGES, IFN-ALPHA HAVE ANTIVIRAL
CC       ACTIVITIES. INTERFERON STIMULATES THE PRODUCTION OF TWO ENZYMES,
CC       A PROTEIN KINASE AND AN OLIGOADENYLATE SYNTHETASE.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: BELONGS TO THE INTERFERON ALPHA, BETA AND DELTA
CC       FAMILY.
CC   ---------------------------------------- ------------ ---- --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------- ------------ ---- --
DR   EMBL; V00545; CAA23806.1; -.
DR   EMBL; K01900; AAA52716.1; -.
DR   EMBL; X03125; CAA26903.1; -.
DR   EMBL; V00550; CAA23811.1; -.
DR   PIR; D23753; IVHUI8.
DR   PIR; A01829; IVHUA4.
DR   HSSP; P01563; 2HIE.
DR   Genew; HGNC:5429; IFNA8.
DR   MIM; 147568; -.
DR   InterPro; IPR000471; Interferon_abd.
DR   Pfam; PF00143; interferon; 1.
DR   PRINTS; PR00266; INTERFERONAB.
DR   ProDom; PD000550; Interferon_abd; 1.
DR   SMART; SM00076; IFabd; 1.
DR   PROSITE; PS00252; INTERFERON_A_B_D; 1.
KW   Cytokine; Antiviral; Multigene family; Signal.
FT   SIGNAL       1     23
FT   CHAIN       24    189       Interferon alpha-8.
FT   DISULFID     24    122       BY SIMILARITY.
FT   DISULFID     92    182       BY SIMILARITY.
FT   CONFLICT      8      8       M -> M (IN REF. 4 AND 5).
FT   CONFLICT     11     11       S/W -> W/S (IN REF. 4 AND 5).
SQ   SEQUENCE   189 AA;  21903 MW;  [illegible] CRC64;

  Query Match             [illegible]  Pred. No. [illegible]
  Best Local Similarity   [illegible]  Gaps [illegible]
  Matches [illegible]  Conservative [illegible]  Mismatches [illegible]  Indels [illegible]

Qy      1 [illegible]
          [illegible]
Db     24 [illegible]

Qy    [illegible]
          [illegible]
Db     84 [illegible]

Qy    120 [illegible]
          [illegible]
Db    144 [illegible]


RESULT 2
INA7_HUMAN
ID   INA7_HUMAN     STANDARD;      [illegible] AA.
AC   P01567; Q14607;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   15-JUN-2000 (Rel. 41, Last annotation update)
DE   Interferon alpha-7 precursor (Interferon alpha-J) (IFN-alpha-
DE   (Interferon alpha-J) (Leu2 ).
GN   IFNA7.
OS   Homo sapiens (human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86047285; PubMed=3057210;
RA   Henco K., Brosius J., Fujisawa A., Fujisawa J.-I., Haynes J.R.,
RA   Hochstadt T., Kawade T., Klett M., Kourilsky P., Maheux R.,
RA   Todokoro K., Waelchli M., Nagata S., Weissmann C.;
RT   "Structural relationship of human interferon alpha genes and
RT   pseudogenes.";
RL   J. Mol. Biol. [illegible] (1985).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE [illegible] PubMed [illegible];
RA   Ullrich A., Gray A., Goeddel [illegible];
RT   "Nucleotide sequence of a partial [illegible] human [illegible]
RT   leukocyte interferon gene cluster.";
RL   J. Mol. Biol. [illegible] (1982).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE [illegible] PubMed [illegible];
RA   Sato S., Weissmann C., Mantei N., Shimizu N.;
RA   Shatzman A.;
RT   "Cloning, expression and [illegible]
RT   interferon alpha hybrids [illegible]
RL   Nucl. Acids Res. [illegible] (1983).
RN   [4]
RP   SEQUENCE OF 24-57.
RX   MEDLINE 98087498; PubMed 9425112;
RA   Nyman T.A., Doelow H., Parkkinen J., Kalkkinen N.;
RT   "Identification of nine interferon-alpha subtypes produced by
RT   virus-induced human peripheral blood leucocytes.";
RL   Biochem. J. [illegible] (1998).
RN   [5]
RP   ABSENCE OF POLYMORPHISM.
RX   MEDLINE [illegible] PubMed [illegible];
```



```
CC       A PROTEIN KINASE AND AN OLIGOADENYLATE SYNTHETASE.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: BELONGS TO THE INTERFERON ALPHA, BETA AND DELTA
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M14541; AAA30950.1; -.
DR   EMBL; A15987; CAA01258.1; -.
DR   PIR; B24912; IVHOA2.
DR   HSSP; P01563; 2HIE.
DR   InterPro; IPR000471; Interferon_abd.
DR   Pfam; PF00143; interferon; 1.
DR   PRINTS; PR00266; INTERFERONAB.
DR   ProDom; PD000550; Interferon_abd; 1.
DR   SMART; SM00076; IFabd; 1.
DR   PROSITE; PS00252; INTERFERON_A_B_D; 1.
KW   Cytokine; Antiviral; Multigene family; Signal.
FT   SIGNAL        1     23
FT   CHAIN        24    184       INTERFERON ALPHA 2.
FT   DISULFID     24    122       BY SIMILARITY.
FT   DISULFID     52    162       BY SIMILARITY.
SQ   SEQUENCE   184 AA;  20877 MW;  [illegible]

  Query Match             74.4%;  Score 633.5;  DB 1:  Length 184;
  Best Local Similarity   76.4%;  Pred. No. 2.3e-51;
  Matches  123;  Conservative   14;  Mismatches   23,  Indels    1,  Gaps    1;

Qy       1 CDLPQTHSLGSRRTLMLLAQMRRISPFSCLKDRHDFGFPQEEFDGNQFQKAPAISVLHELIQQ 64
           |||| ||||||: | |||| |||||| |||||||:||||||| | ||||||: : | :|.
Db      24 CPLPHTHSLGNTRVLMLIGQMRKISPFSCLKDRNDFRFPEEVFDGNQFQKPQAISVLHEMIQQ 83

Qy      65 IQQIFNLFSTKDSSAAWDETLLDKFYIELYQQLNDLEACVIQGVGVTETPLMKEDSILAV 124
           |||||:|||| ||||||||:|||| || ||||| :|||| : ||| |||||| :|?:|:|
Db      84 IQQIFHLFSTDGSSAAWDESLLDKLYTGLYQQLEDLEACLSQEVGVEETPLMNEDSILAV 143

Qy     125 KKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTNLQES 164
           |:|||| |||:|||||||||||:|||||| || ||||||:|
Db     144 KRYFQRIALYLQEKKYSPCAWEIVRAEIMRSFSSSTNLQQS 184


RESULT 15
INAI_HORSE
ID   INAI_HORSE     STANDARD;      PRT;   184 AA.
AC   P05003;
DT   13-AUG-1987 (Rel. 05, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Interferon alpha-1 precursor.
OS   Equus caballus (Horse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX   NCBI_TaxID=9796;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87053170; PubMed=3022999;
RA   Himmler A., Hauptmann R., Adolf G.R., Swetly P.;
RT   "Molecular cloning and expression in Escherichia coli of equine type
RT   I interferons.";
RL   DNA 5:345-356(1986).
CC   -!- FUNCTION: PRODUCED BY MACROPHAGES, IFN-ALPHA HAVE ANTIVIRAL
CC       ACTIVITIES. INTERFERON STIMULATES THE PRODUCTION OF TWO ENZYMES:
CC       A PROTEIN KINASE AND AN OLIGOADENYLATE SYNTHETASE.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: BELONGS TO THE INTERFERON ALPHA, BETA AND DELTA
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M14540; AAA30949.1; -.
DR   PIR; A15912; IVHOA1.
DR   HSSP; P01563; 2HIE.
DR   InterPro; IPR000471; Interferon_abd.
DR   Pfam; PF00143; interferon; 1.
DR   PRINTS; PR00266; INTERFERONAB.
DR   ProDom; PD000550; Interferon_abd; 1.
DR   SMART; SM00076; IFabd; 1.
DR   PROSITE; PS00252; INTERFERON_A_B_D; 1.
KW   Cytokine; Antiviral; Multigene family; Signal.
FT   SIGNAL        1     23
FT   CHAIN        24    184       INTERFERON ALPHA 1.
FT   DISULFID     24    122       BY SIMILARITY.
FT   DISULFID     52    162       BY SIMILARITY.
SQ   SEQUENCE   184 AA;  20877 MW;  [illegible]

  Query Match             74.4%;  Score 633.5;  DB 1:  Length 184;
  Best Local Similarity   76.4%;  Pred. No. 2.3e-51;
  Matches  123;  Conservative   14;  Mismatches   23,  Indels    1,  Gaps    1;

Qy       1 CDLPQTHSLGSRRTLMLLAQMRRISPFSCLKDRHDFGFPQEEFDGNQFQKAPAISVLHELIQQ 64
           |||| ||||||: | |||| |||||| |||||||:||||||| | ||||||: : | :|.
Db      24 CPLPHTHSLGNTRVLMLIGQMRKISPFSCLKDRNDFRFPEEVFDGNQFQKPQAISVLHEMIQQ 83

Qy      65 IQQIFNLFSTKDSSAAWDETLLDKFYIELYQQLNDLEACVIQGVGVTETPLMKEDSILAV 124
           |||||:|||| ||||||||:|||| || ||||| :|||| : ||| |||||| :|?:|:|
Db      84 IQQIFHLFSTDGSSAAWDESLLDKLYTGLYQQLEDLEACLSQEVGVEETPLMNEDSILAV 143

Qy     125 KKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTNLQES 164
           |:|||| |||:|||||||||||:|||||| || ||||||:|
Db     144 KRYFQRIALYLQEKKYSPCAWEIVRAEIMRSFSSSTNLQQS 184


Search completed: January 17, 2003, 09:14:27
Total time: 13.04 seconds
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: January 17, 2003, 06:11:19 ; Search Time 14.6129 Seconds
(without alignments)
699.471 Million cell updates/sec

Title: US-09-915-873-5
Perfect score: 851
Sequence: 1 CDLPQTHSLGSRRTLMLLAQ.....EIMRSFSLSTNLQESLRSKE 166

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_2:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	741.5	86.0	189	6	Q95J78	Q95j78 sapajus
2	706.5	83.0	189	6	Q95J77	Q95j77 saguinus
3	692.5	81.4	181	4	Q14608	Q14608 homo sapien
4	688.5	80.9	166	4	Q9UM13	Q9um13 homo sapien
5	688.5	80.9	189	4	Q14605	Q14605 homo sapien
6	685.5	80.6	189	4	Q14619	Q14619 homo sapien
7	677.5	79.6	166	4	Q8WYC9	Q8wyc9 homo sapien
8	529.5	62.2	190	11	Q61718	Q61718 mus musculu
9	524.5	61.6	190	11	Q61719	Q61719 mus musculu
10	523	61.5	190	11	Q99N22	Q99n22 marmota mon
11	517	60.8	190	11	Q8K5J5	Q8k5j5 marmota mon
12	508.5	59.8	190	11	Q64138	Q64138 cricetulus
13	507	59.6	201	11	Q8K5K9	Q8k5k9 marmota mon
14	504.5	59.3	190	11	Q61716	Q61716 mus musculu
15	502.5	59.0	166	6	P79228	P79228 sabaeus
16	501	58.9	190	11	Q8K5J4	Q8k5j4 marmota mon

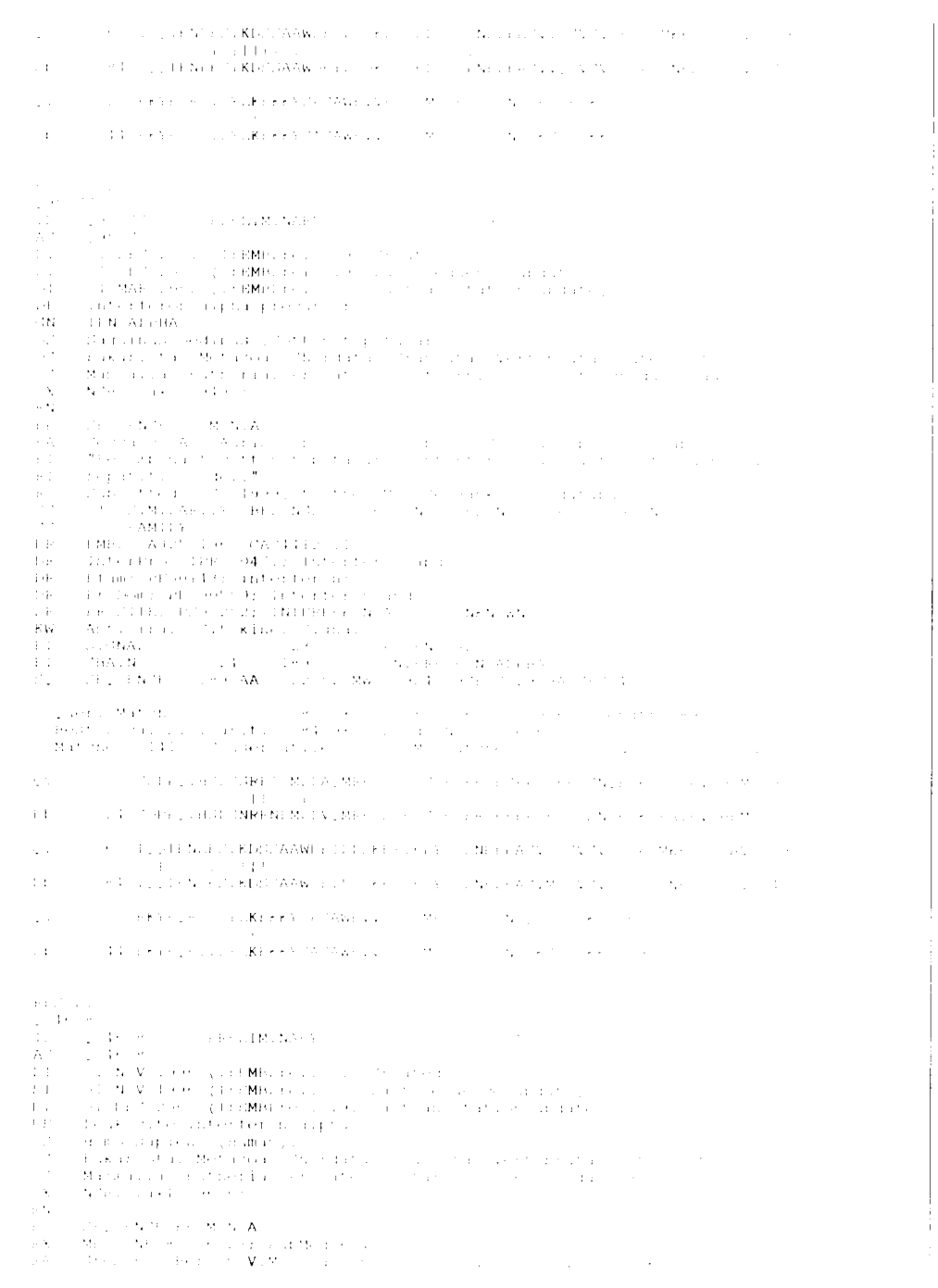
17	500	58.8	190	11	Q8K5J3	Q8k5j3 marmota mon
18	494.5	58.1	171	5	Q15323	Q15323 homo sapien
19	492	57.8	190	11	Q64139	Q64139 cricetulus
20	490	57.6	197	11	Q8K5L7	Q8k5l7 marmota mon
21	490	57.6	197	11	Q8K5L7	Q8k5l7 marmota mon
22	487	57.2	197	11	Q9WVH5	Q9wvh5 mus musculu
23	486	57.1	190	11	Q8K5J1	Q8k5j1 marmota mon
24	485	57.0	197	11	Q9WVH4	Q9wvh4 mus musculu
25	485	57.0	197	11	Q9WVH6	Q9wvh6 mus musculu
26	479	56.3	190	11	Q64137	Q64137 cricetulus
27	472.5	55.5	195	4	Q96AA7	Q96aa7 homo sapien
28	472.5	55.5	195	4	Q96AA8	Q96aa8 homo sapien
29	472.5	55.5	195	4	Q96AA6	Q96aa6 homo sapien
30	472.5	55.5	195	4	Q96AA5	Q96aa5 homo sapien
31	471.5	55.4	195	4	Q96AA9	Q96aa9 homo sapien
32	469.5	55.2	195	4	Q96AB0	Q96ab0 homo sapien
33	449	52.8	195	4	Q96AB1	Q96ab1 homo sapien
34	447.5	52.6	195	4	Q7Z5Z8	Q7z5z8 homo sapien
35	442.5	52.0	195	4	Q96AB2	Q96ab2 homo sapien
36	440	51.7	179	4	Q53QM1	Q53qm1 homo sapien
37	436.5	51.3	195	4	Q9NYB9	Q9nyb9 homo sapien
38	431.5	50.7	195	4	Q96AI7	Q96ai7 homo sapien
39	425.5	50.0	195	4	Q7Z5Z7	Q7z5z7 homo sapien
40	424	49.8	195	4	Q96AI6	Q96ai6 homo sapien
41	422.5	49.6	195	11	Q9JLI2	Q9jli2 mus musculu
42	422.5	49.6	195	11	Q9JLW5	Q9jlw5 mus musculu
43	416.5	48.9	168	4	Q9NZU8	Q9nzu8 homo sapien
44	408	48.0	201	11	Q9R0T3	Q9r0t3 mus musculu
45	402	47.3	95	1	Q11431	Q11431 equus caba

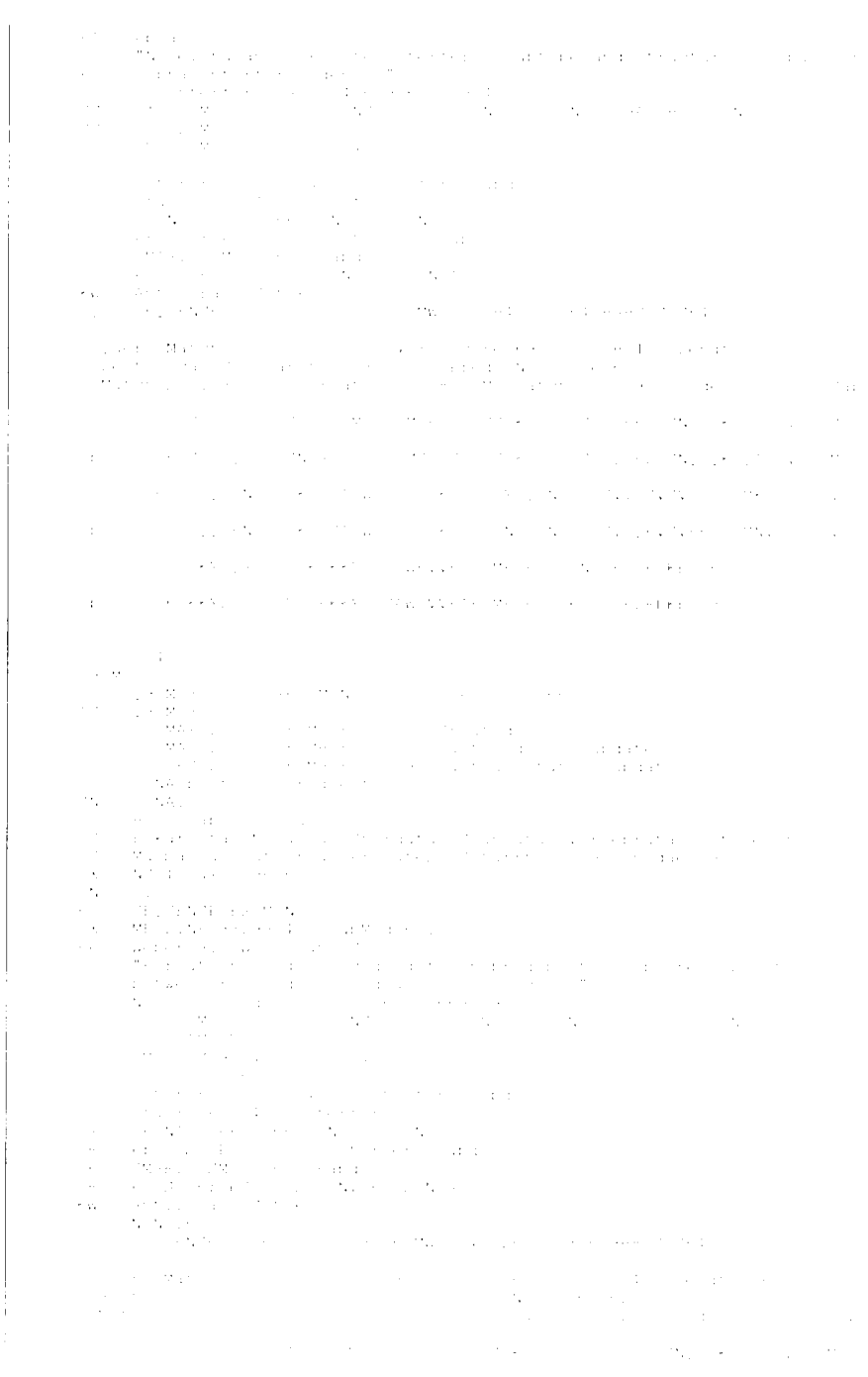
ALIGNMENTS

RESULT 1
Q95J78
ID Q95J78 PRELIMINARY; 189 AA.
AC Q95J78;
DT 01-DEC-2001 (TrEMBLrel. 14, Created)
DT 01-DEC-2001 (TrEMBLrel. 14, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 15, Last annotation update)
DE Interferon alpha precursor.
GN IFN-ALPHA.
OS Saguinus oedipus (Cotton-top tamarin).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euprimates; Haplorrhini; Platyrrhini; Callitrichidae;
OC Saguinus.
OX NCBI_TaxID=9490;
RN [1]
RP SEQUENCE FROM N.A.
RA Cecconi A., Antisichella A., Chiaberti S., Palombo F., Tonlorenzi R.;
RT "Recombinant cotton-top tamarin interferon-alpha: a new tool for
RT hepatitis models.";
RL Submitted (01-DEC-2001) to EMBL/GenBank/DDBJ databases.
CC -!- SIMILARITY: BELONGS TO THE INTERFERON ALPHA, BETA AND OMEGA
CC FAMILY.
DR EMBL; AJ250195; CAC44124.1; -.
DR InterPro; IPR000471; Interferon_a/b.
DR Pfam; PF00143; interferon; 1.
DR ProDom; PD000555; Interferon_a/b; 1.
DR PROSITE; PS00252; INTERFERON_A_B_OMEGA; 1.
KW Antiviral; Cytokine; Secreted.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 189 INTERFERON.
SQ SEQUENCE 189 AA; 21937 MW; 9A71E1206 CRC64;

Query Match 86.0%; Score 741.5; DB 6; Length 189;
Best Local Similarity 87.4%; Pred. No. 1.1e-65;
Matches 145; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

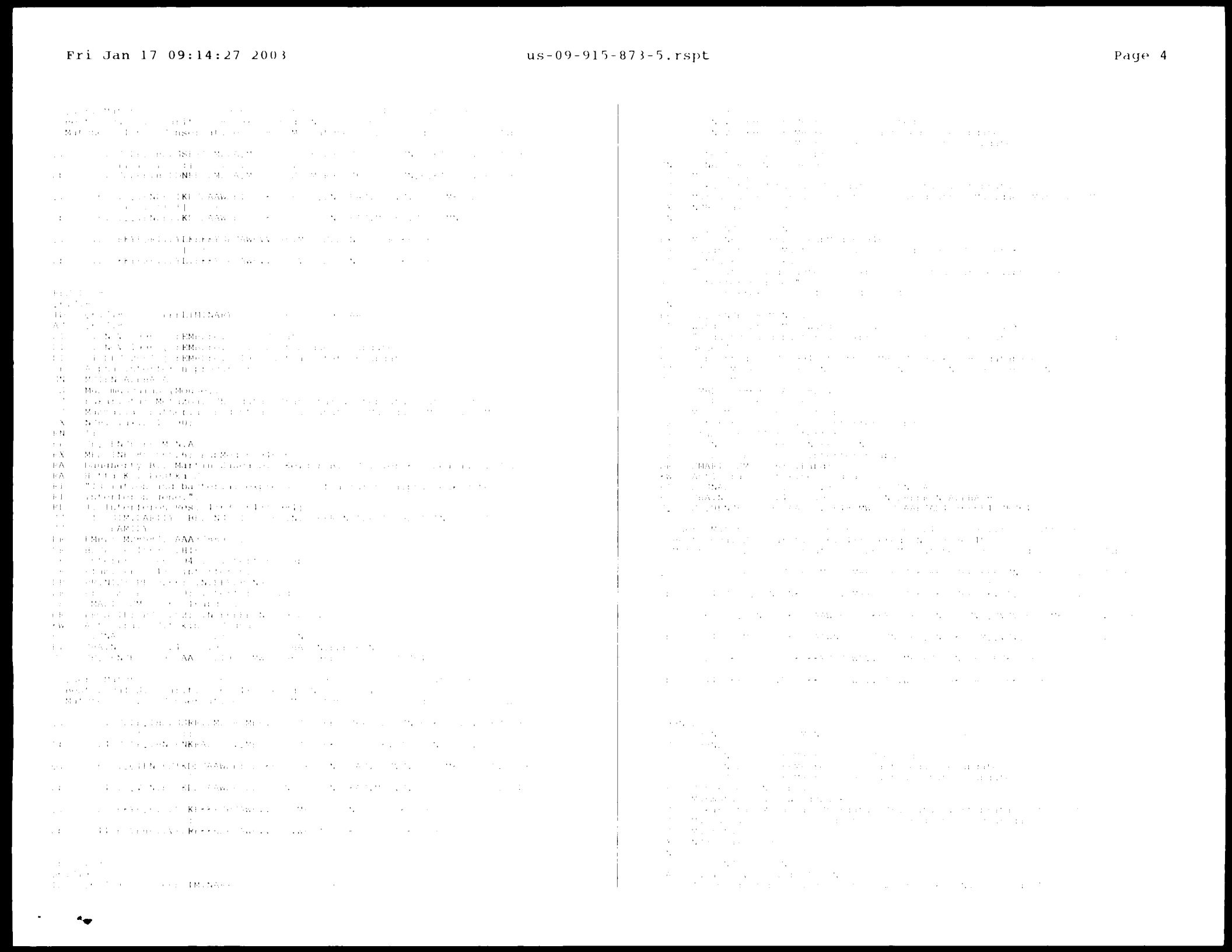
Qy 1 CDLPQTHSLGSRRTLMLLAQMRRISPFSCLKDRHDFGFPQEEFDGNQFQKAQAISVLHEMIQQTFNLFSTKDSSAAWDETLLDKFYTELYQQLNDLEACVIQGVGVTETPLMKEDSILAVRKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTNLQESLRSKE 166
||||||||||:|||||||:|:|||| |:| |||| || |||||| |||||| |||| | |
Db 24 CDLPQTHSLGSRRTLMLLAQMRRISPFSCLKDRHDFGFPQEEFDGNQFQKAQAISVLHEMIQQTFNLFSTKDSSAAWDETLLDKFYTELYQQLNDLEACVIQGVGVTETPLMKEDSILAVRKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTNLQESLRSKE 189



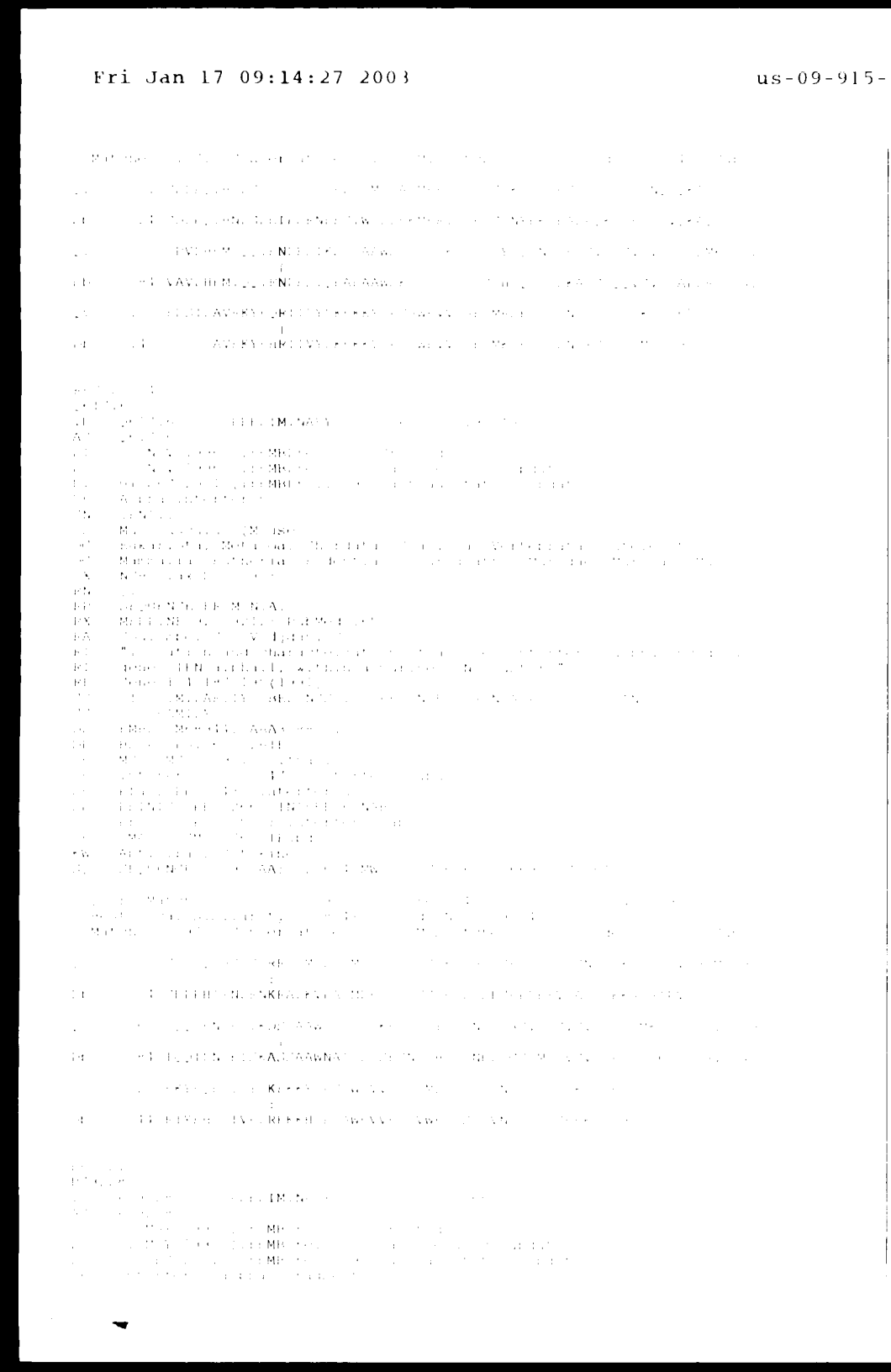


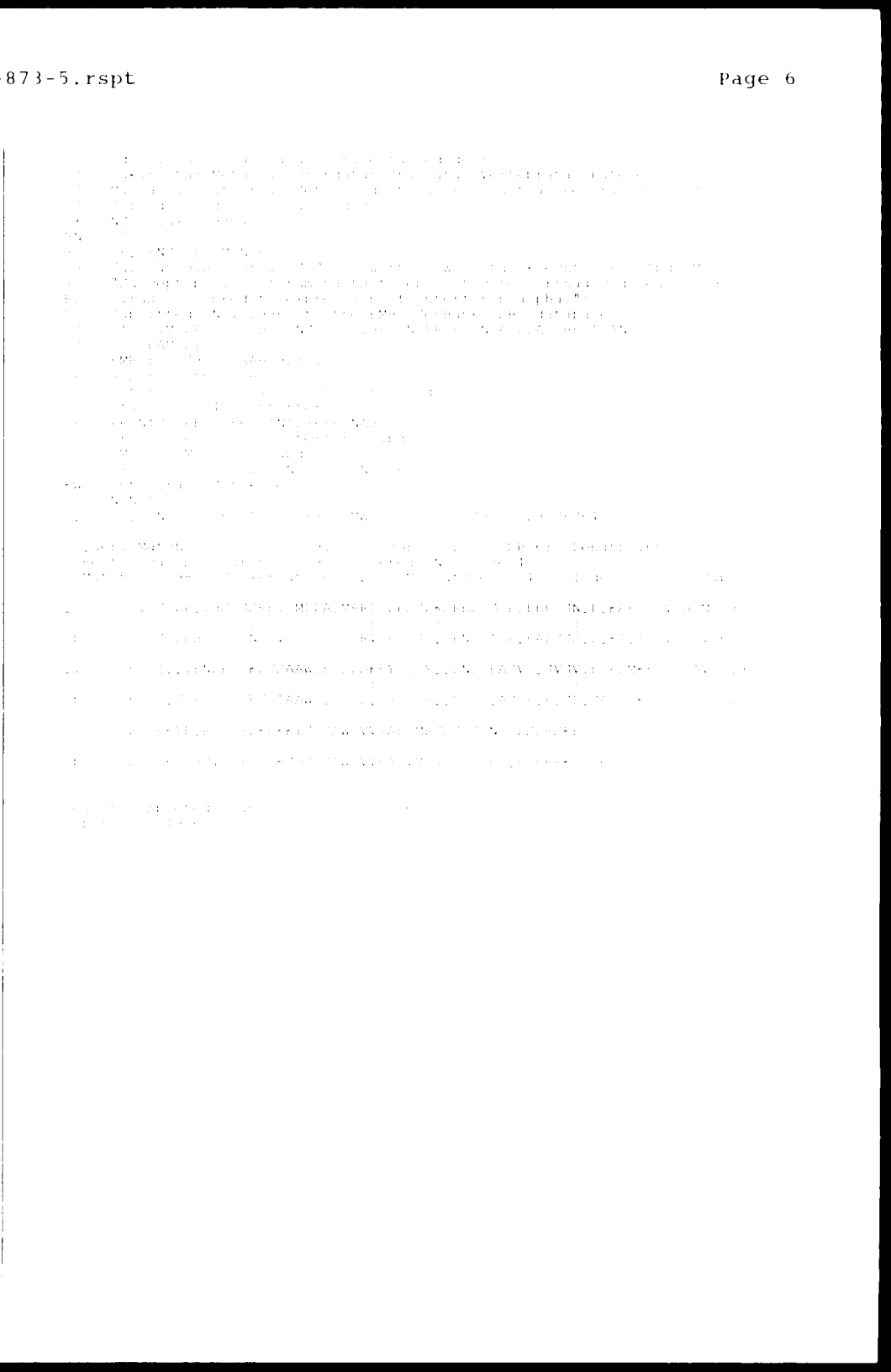
```
Db      1 CDLPETHSLDNRRTLMLLAQMSRISPSSCLMDRHDFGFPQEEFDGNQFQKAPAISVLHEL 60

Qy     60 IQQIFNLFSTKDSSAAWDETLLDKFYIELYQQLNDLEACVMQEERVGETPLMNADSILAV 119
          ||||||||:|||||||||| |||| ||||||||||||||:|   | |||||  ||||||
Db     61 IQQIFNLFTTKDSSAAWDEDLLDKFCIELYQQLNDLEACVMQEERVGETPLMNADSILAV 120

Qy    120 KKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTNLQESLRSKE 165
          :|||:|||||| ||||| ||||||||||||| |||||||| || ||
Db    121 KKYFHRISVYLTEKKYSPCAWEVVRAFIMRSLSFSTNLQKRLKRKE 166


RESULT 5
Q14605
ID   Q14605      PRELIMINARY;      PRT:   189 AA.
AC   Q14605;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last Sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   Interferon-alpha 13 precursor.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Rostoks N.;
RL   Submitted (JAN-1994) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86037205; PubMed=4057246;
RA   Henco K., Brosius J., Fujisawa A., Fujisawa J.I., Haynes J.R.,
RA   Hochstadt J., Kovacic S., Pasek M., Schultz K.A., Schmid J.,
RA   Todokoro K., Waelchli M., Nagata S., Weissmann C.;
RT   "Structural relationship of human interferon alpha genes and
RT   pseudogenes.";
RL   J. Mol. Biol. 185:227-260(1985).
CC   -!- SIMILARITY: BELONGS TO THE INTERFERON ALPHA, BETA AND DELTA
CC       FAMILY.
DR   EMBL; X75924; CAA53538.1; -.
DR   EMBL; A07164; CAA00632.1; -.
DR   HSSP; P01563; 2HIE.
DR   InterPro; IPR000471; Interferon_abd.
DR   Pfam; PF00143; Interferon; 1.
DR   PRINIS; PR00266; INTERFERONAB.
DR   ProDom; PD000550; Interferon_abd; 1.
DR   SMART; SM00076; IFabd; 1.
DR   PROSITE; PS00252; INTERFERON_A_B_D; 1.
KW   Antiviral; Cytokine; Signal.
FT   SIGNAL        1     23       POTENTIAL.
FT   CHAIN        24    189       POTENTIAL.
SQ   SEQUENCE   189 AA;  21697 MW;  442F88B754298798 CRC64;

  Query Match             80.9%;  Score 698.5;  DB 4;  Length 189;
  Best Local Similarity   83.1%;  Pred. No. 1e-61;
  Matches  138;  Conservative    7;  Mismatches   20;  Indels    1;  Gaps    1;

Qy      1 CDLPQTHSLGSRRTLMLLAQMRRISLFSCLKDRHDFGFPQEEFDGNQFQKAEAISVLHEM 60
          |||:|||| :|||||||||| ||| ||| ||||||||||||| |||||||  | ||||
Db     24 CDLPETHSLDNRRTLMLLAQMSRISPSSCLMDRHDFGFPQEEFDNQFQKAPAISVLHEL 83

Qy     60 IQQIFNLFSTKDSSAAWDETLLDKFYTELYQQLNDLEACVIQGVGVTETPLMKEDSILAV 119
          ||||||||:|||||||||| |||| ||||||||||||:|||  | ||||  ||||||
Db     84 IQQIFNLFTTKDSSAAWDEDLLDKFCTELYQQLNDLEACVMQEERVGETPLMNADSILAV 143

Qy    120 KKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTNLQESLRSKE 165
          :|||:|||||| ||||| ||||||||||||| |||||||| || ||
Db    144 KKYFHRISLYLTEKKYSPCAWEVVRAEIMRSLSLSTNLQERLKRKE 189


RESULT 6
Q14618
ID   Q14618      PRELIMINARY;      PRT:   189 AA.
```

```
AC   Q14605;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last Sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   Interferon alpha-N3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Chen K.Y., Bonnin V.M., Isaguliants A.Y., Agasti S.Ch., Vishnevsky A.,
RA   Yansoone I.V., Dishler A.V., Pudova N.V., Smirnodilatova A.A.,
RA   Isaeva V.I., Stepanov A.N., Isaeva I.Y., Medvedis Y.A., Ch. V.I.,
RA   Karpas V.M., Isaeva V.A., Sokil O.I.;
RT   "A new type of leukocyte interferon.";
RL   Dokl. Biochem. 369:91-95(1999).
CC   -!- SIMILARITY: BELONGS TO THE INTERFERON ALPHA, BETA AND DELTA
CC       FAMILY.
DR   EMBL; X99149; CAA67470.1; -.
DR   HSSP; P01563; 2HIE.
DR   InterPro; IPR000471; Interferon_abd.
DR   Pfam; PF00143; Interferon; 1.
DR   PRINIS; PR00266; INTERFERONAB.
DR   ProDom; PD000550; Interferon_abd; 1.
DR   SMART; SM00076; IFabd; 1.
DR   PROSITE; PS00252; INTERFERON_A_B_D; 1.
KW   Antiviral; Cytokine.
SQ   SEQUENCE   189 AA;  21810 MW;  [illegible] CRC64;

  Query Match             [illegible]
  Best Local Similarity   [illegible]
  Matches  [illegible]

Qy      1 [illegible]
          |||||||||||:|| :|||||  |||  | ||||||||||| ||  | |||||  ||
Db     24 [illegible]

Qy     60 IQQIFNLFSTKDSSAAW[illegible]
          ||  |||||||||||:|  |  |  |  |  |  |  |  |  |  |  |  |  |
Db     84 IQQIFNLFTTKDSSAAWDEDLLDKFCTELYQQLNDLEACVMQEERVGETPLMNADSILAV 143

Qy    120 KKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTNLQESLRSKE 165
          :|||:|||||| ||||| ||||||||||||| |||||||| || ||
Db    144 KKYFHRISLYLTEKKYSPCAWEVVRAEIMRSLSLSTNLQERLKRKE 189


RESULT 7
Q8WZ68
ID   Q8WZ68      PRELIMINARY;      PRT:   [illegible] AA.
AC   Q8WZ68;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last Sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Interferon alpha 1b (Fragment).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; [illegible]
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Chen H.H., Yu X.B.;
RL   Submitted ([illegible]) to the EMBL/GenBank/DDBJ databases.
RL   EMBL; [illegible]; AAL[illegible]; -.
DR   InterPro; IPR000471; Interferon_abd.
DR   Pfam; PF00143; Interferon; 1.
DR   PRINIS; PR00266; INTERFERONAB.
DR   ProDom; PD000550; Interferon_abd; 1.
DR   SMART; SM00076; IFabd; 1.
DR   PROSITE; PS00252; INTERFERON_A_B_D; 1.
FT   NON_TER       1       1
SQ   SEQUENCE   166 AA;  [illegible] MW;  [illegible] CRC64;
```



```
RL   Submitted (FEB 2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO THE INTERFERON ALPHA, BETA AND DELTA
CC       FAMILY.
DR   EMBL; AF231139; AAK14944.1; -.
DR   HSSP; P01563; 2HIE.
DR   InterPro; IPR000471; Interferon_abd.
DR   Pfam; PF00143; interferon; 1.
DR   PRINTS; PR00266; INTERFERONAB.
DR   ProDom; PD000550; Interferon_abd; 1.
DR   SMART; SM00076; IFabd; 1.
DR   PROSITE; PS00252; INTERFERON_A_B_D; 1.
KW   Antiviral; Cytokine.
SQ   SEQUENCE   190 AA;  21810 MW;  1875B940CA56834E CRC64;

  Query Match             61.5%;  Score 523;  DB 11;  Length 190;
  Best Local Similarity   63.0%;  Pred. No. 5.3e-45;
  Matches  109;  Conservative   16;  Mismatches   84;  Indels   14;  Gaps    4;

Qy       1 CDLPQTHSLG-------SRRTLMLLAQMRRISLFSCLKDRHDFGFPQEEF-GNQFQKAET 52
           ||||| |:||          | |:| :|||  |||| | || |:|:  | | |||
Db      24 CDLPQTHNLGLETPDENEPSTWMVLTKMRFIFLFSTLEYSPEDTAFTRELLQELYVLKALA 83

Qy      53 IPVLHEMIQQIFNLFSTKDSSAAWNETLLEKFYIELYQQLNDLEACVIQGVGVTETPLMK 112
           : ||||| |||||||||:::|||||:|||| | | |:|||||:|||:| : ||| | ||
Db      84 VAVLREMTQQIFNLFSTQDASAAWDKTLLDKFLTELYQQLNDLEACVIQQVGVEEAPLM  142

Qy     113 EDSILAVRKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTNLQESLRSKE 165
                ||||| |||:|||||| |||||||||||||||| | |:   ||| :
Db     143      AVRKYFHRITVYLKEKKYSPCAWEVVRAEIMKSFSLSTNLQKSLKSKE 190


RESULT 11
Q8R5J5
ID   Q8R5J5      PRELIMINARY;      PRT;   190 AA.
AC   Q8R5J5;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Interferon alpha 3.
OS   Marmota monax (Woodchuck).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciuromorpha; Sciuridae; Sciurinae;
OC   Marmota.
OX   NCBI_TaxID=9995,
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Berraondo P., Garcia R., Blanco B., Vales A., Ruiz J., Prieto J.;
RT   "Woodchuck interferon alfa: Cloning and evaluation of hepatic
RT   antiviral activity.";
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF425727; AAL26915.1; -.
SQ   SEQUENCE   190 AA;  21749 MW;  394ABCF0E9E71B59 CRC64;

  Query Match             60.8%;  Score 517;  DB 11;  Length 190;
  Best Local Similarity   62.4%;  Pred. No. 2.1e-44;
  Matches  108;  Conservative   16;  Mismatches   35;  Indels   14;  Gaps    4;

Qy       1 CDLPQTHSLG-------SRRTLMLLAQMRRISLFSCLKDRHDFGFPQEEF-GNQFQKAET 52
           ||||| |:||          | |:| :|||  |||| | || |:|:  | | |||
Db      24 CDLPQTHNLGLETPDENEPSTWMVLTKMRFIFLFSTLEYSPEDTALTRELLQELYVLKALA 83

Qy      53 IPVLHEMIQQIFNLFSTKDSSAAWNETLLEKFYIELYQQLNDLEACVIQGVGVTETPLMK 112
           : ||||| |||||||||:::|||||:|||| | | |:|||||:|||:| : ||| | ||
Db      84 VAVLREMTQQIFNLFSTQDASAAWDKTLLDKFLTELYQQLNDLEACVIQQVGVEEAPLM  142

Qy     113 EDSILAVRKYFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTNLQESLRSKE 165
                ||||| |||:|||||| |||||||||||||||| | |:   ||| :
Db     143      AVRKYFHRITVYLKEKKYSPCAWEVVRAEIMKSFSLSTNLQKSLKSKE 190


RESULT 12
Q64138
ID   Q64138      PRELIMINARY;      PRT;   190 AA.
AC   Q64138;
DT   01-NOV 1996 (TrEMBLrel. 01, Created)
DT   01-NOV 1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC 2001 (TrEMBLrel. 19, Last annotation update)
DE   Interferon-alpha 1.
GN   IFA-1.
OS   Cricetinae (hamsters).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciuromorpha; Muridae;
OX   NCBI_TaxID=10029;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 95311634; PubMed=7600693;
RA   Provvedi R.A., Nikamura N., Vilcek J., Pestka S.;
RA   Unpublished.;
RT   "Isolation, properties and chromosomal localization of the human
RT   linked hamster interferon-alpha genes.";
RL   Gene 156:299-300(1995).
CC   -!- SIMILARITY: BELONGS TO THE INTERFERON ALPHA, BETA AND DELTA
CC       FAMILY.
DR   EMBL; X78750; CAA55444.1; -.
DR   HSSP; P01563; 2HIE.
DR   InterPro; IPR000471; Interferon_abd.
DR   Pfam; PF00143; interferon; 1.
DR   PRINTS; PR00266; INTERFERONAB.
DR   ProDom; PD000550; Interferon_abd; 1.
DR   SMART; SM00076; IFabd; 1.
DR   PROSITE; PS00252; INTERFERON_A_B_D; 1.
KW   Antiviral; Cytokine.
SQ   SEQUENCE   190 AA;  21618 MW;  B5E14F0C1A_6A9CA1 CRC64;

  Query Match             59.8%;  Score 508;  DB 11;  Length 190;
  Best Local Similarity   59.6%;  Pred. No. 1.1e-43;
  Matches   99;  Conservative   20;  Mismatches   40;  Indels    4;  Gaps    1;

Qy       1 CDLPQTHSLGSRRTLMLLAQMRRISLFSCLKDRHDFGFPQEEFGNQFQKAETIPVLHEMI 60
           |:|| |:|:  | | |||  |||| | || |:|:  | | |||  |||| | |||:|||
Db      24 CDLPQTHNLETPDENEPSTWMVLTKMRFIFLFSTLEYSPEDTALTRELLQELYVLKALAS 83

Qy      61 QQIFNLFSTKDSSAAWNETLLEKFYIELYQQLNDLEACVIQGVGVTETPLMKEDSILAVRK 120
           | | |||:|||:|| |||:|  |  | |:|:  |:|  | |  |||:|||:|| |||:|||
Db      84 QQVFNLFSTEDSSAAWEQSLLEEFYTELYQQLNDLKACVMQEVGVMETPLMNVDSILAVKK 143

Qy     121 FFQRITLYLKEKKYSPCAWEVVRAEIMRSFSLSTNLQESLRSKE 165
           || | |||:|||:|| |||:|  |  | |:|:  |:|  | |  ||
Db     144 KFYYIHGRLTHYLKEKEYSPCAWEVVRVEIMRSLSFSTNLQKRLRRKD 190


RESULT 13
Q9R5K9
ID   Q9R5K9      PRELIMINARY;      PRT;   190 AA.
AC   Q9R5K9;
DT   01-JUN 2002 (TrEMBLrel. 21, Created)
DT   01-JUN 2002 (TrEMBLrel. 21, Last sequence update)
DT   01 JUN 2002 (TrEMBLrel. 21, Last annotation update)
DE   Interferon alpha 4.
GN   IFNA4.
OS   Marmota monax (Woodchuck).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciuromorpha; Sciuridae; Sciurinae;
OC   Marmota.
OX   NCBI_TaxID=9995;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Berraondo P., Garcia R., Blanco B., Vales A., Ruiz J., Prieto J.;
RT   "Woodchuck interferon alpha subtypes: cDNA cloning and evaluation of
RT   activity.";
RL   Submitted (FEB 2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF338277; AAL71997.2; -.
SQ   SEQUENCE   190 AA;  21763 MW;  43B  [illegible]  CRC64;

  Query Match             61.8%;  Score 525;  DB 11;  Length 190;
  Best Local Similarity   61.8%;  Pred. No. 2.1e-45;
```





SeqCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: January 17, 2003, 05:08:15 ; Search time 50 Seconds
(without alignments)
3033.909 Million cell updates/sec

Title: US-09-915-873-1
Perfect score: 554
Sequence: 1 gatcaaqtqcaaaaqtcaqc.....ctqqaatccatcaaactqaqa 554

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfile1.seq:*

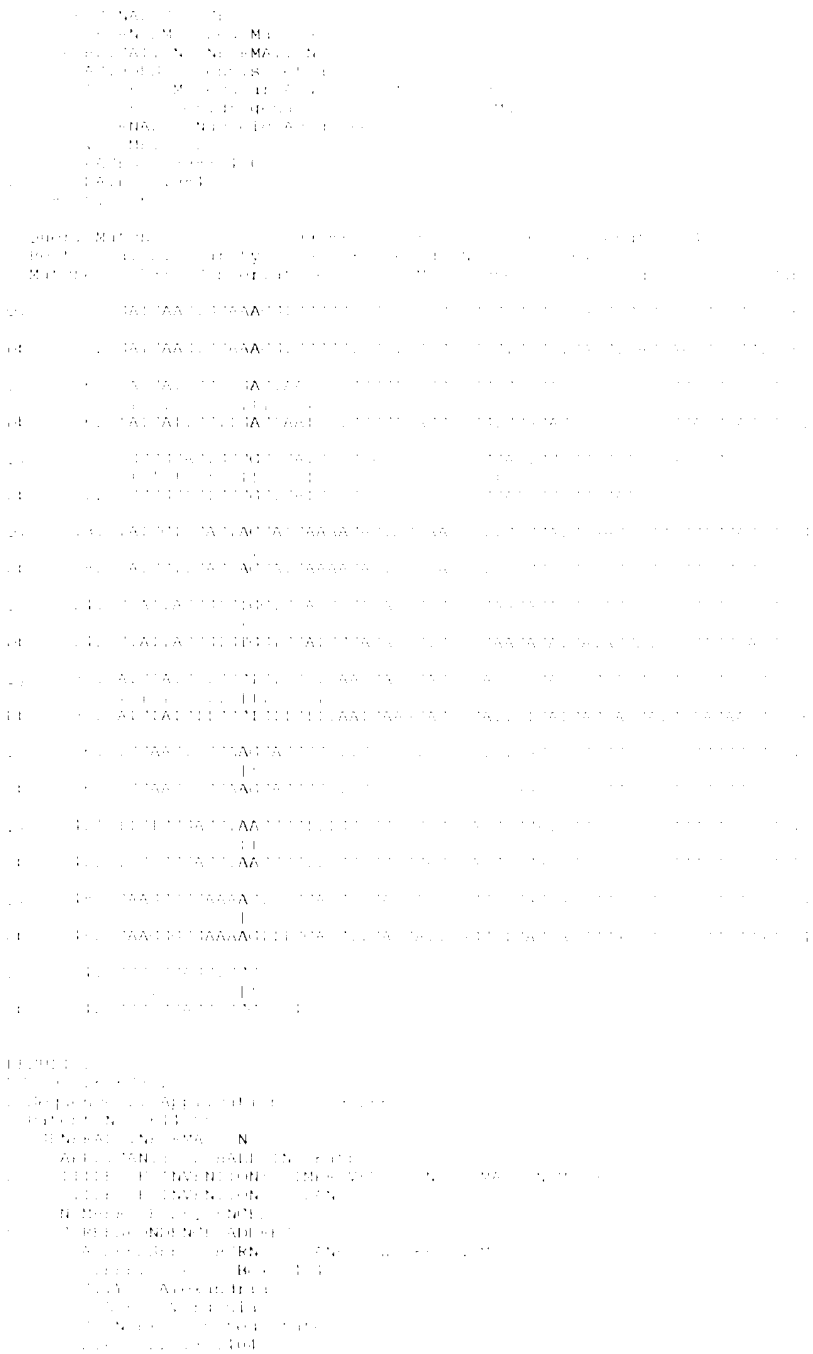
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

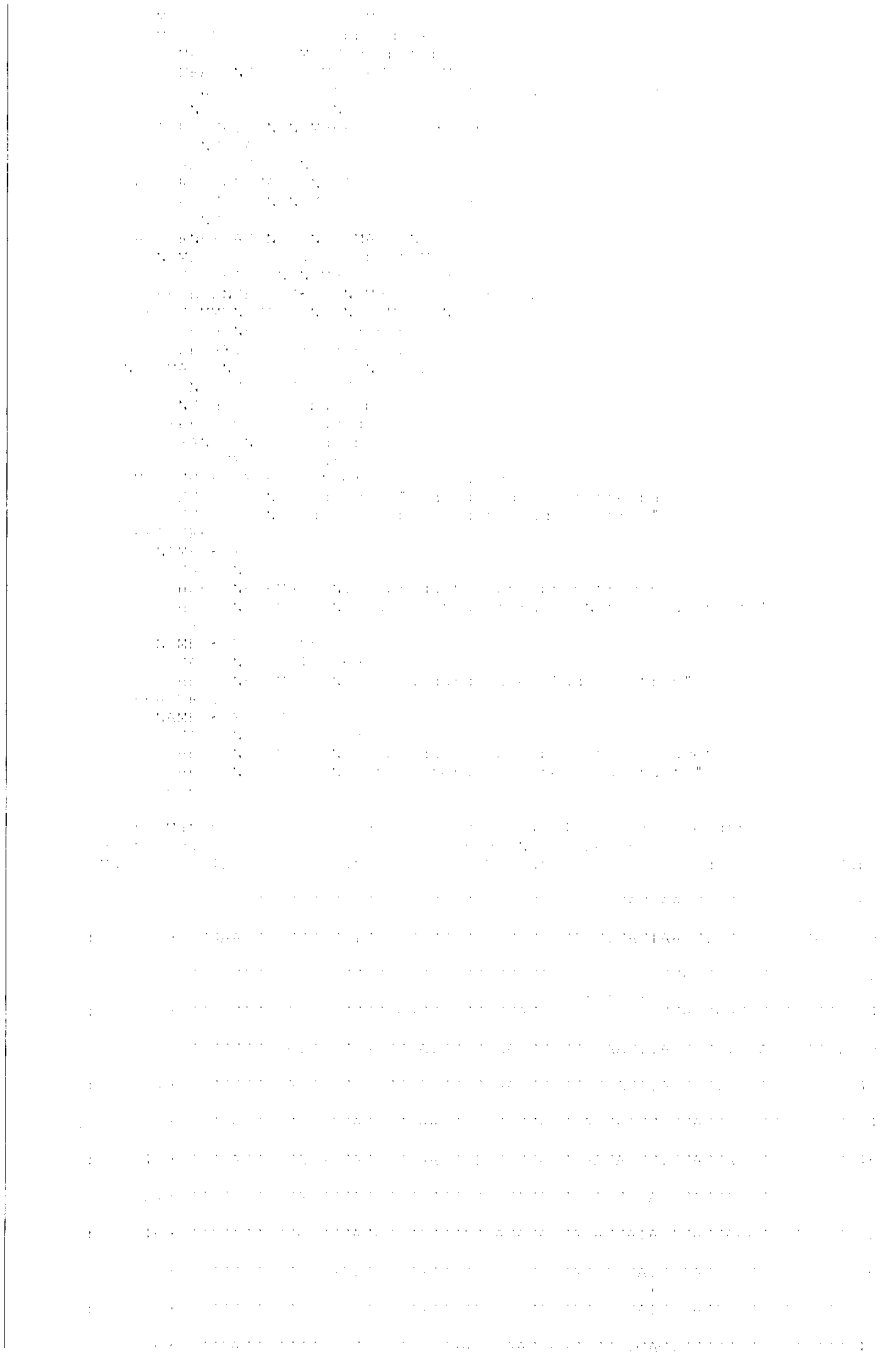
SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	553	99.8	554	1	US-08-051-006-5	Sequence 5, Appli
2	541.4	97.7	1086	4	US-09-026-673-2	Sequence 2, Appli
3	541.4	97.7	1086	4	US-09-480-142-2	Sequence 2, Appli
4	540	97.5	5033	1	US-08-038-768A-1	Sequence 1, Appli
5	119	21.5	5897	4	US-09-097-319A-26	Sequence 26, Appl
6	119	21.5	9335	4	US-09-097-319A-19	Sequence 19, Appl
7	118	21.3	219	4	US-08-729-601A-47	Sequence 47, Appl
8	118	21.3	380	4	US-08-729-601A-44	Sequence 44, Appl
9	118	21.3	1030	1	US-07-936-163-46	Sequence 46, Appl
10	118	21.3	1030	4	US-08-729-601A-43	Sequence 43, Appl
11	118	21.3	10160	4	US-09-097-319A-8	Sequence 8, Appli
12	118	21.3	11784	4	US-09-097-319A-9	Sequence 9, Appli
13	118	21.3	11991	4	US-09-097-319A-10	Sequence 10, Appl
14	43	7.8	7218	1	US-08-232-464-14	Sequence 14, Appl
c 15	34.2	6.2	2238	4	US-08-939-366-5	Sequence 5, Appli
16	34	6.1	1647	1	US-08-066-299-12	Sequence 12, Appl
17	34	6.1	1647	1	US-08-383-756-7	Sequence 7, Appli
18	34	6.1	1647	1	US-08-265-047-3	Sequence 3, Appli
19	34	6.1	1647	2	US-08-460-868-7	Sequence 7, Appli
c 20	32.6	5.9	1530	4	US-09-134-001C-159	Sequence 159, App
c 21	31.8	5.7	3209	1	US-08-105-483-220	Sequence 220, App
c 22	31.8	5.7	3209	1	US-08-220-151-62	Sequence 62, Appl
c 23	31.8	5.7	3209	1	US-08-413-119-62	Sequence 62, Appl
c 24	31.8	5.7	3209	1	US-08-224-391-98	Sequence 98, Appl
c 25	31.8	5.7	3209	1	US-08-484-304-98	Sequence 98, Appl
c 26	31.8	5.7	3209	1	US-08-224-657-99	Sequence 99, Appl
c 27	31.8	5.7	3209	1	US-08-709-209-220	Sequence 220, App
c 28	31.8	5.7	4209	1	US-08-510-[illegible]	Sequence [illegible]
c 29	31.8	5.7	4209	1	US-08-[illegible]	Sequence [illegible]
c 30	31.8	5.7	4209	1	US-08-[illegible]	Sequence [illegible]
c 31	31.8	5.7	4209	1	US-08-[illegible]	Sequence [illegible]
c 32	31.8	5.7	4209	1	US-08-[illegible]	Sequence [illegible]
c 33	31.8	5.7	4209	1	US-08-[illegible]	Sequence [illegible]
c 34	31.8	5.7	4209	1	US-08-[illegible]	Sequence [illegible]
c 35	31.8	5.7	4209	1	US-08-[illegible]	Sequence [illegible]
c 36	31.8	5.7	4209	1	US-08-[illegible]	Sequence [illegible]
c 37	31.8	5.7	4209	1	US-08-[illegible]	Sequence [illegible]
c 38	31.8	5.7	4209	1	US-08-[illegible]	Sequence [illegible]
c 39	31.8	5.7	4209	1	US-08-[illegible]	Sequence [illegible]
c 40	31.8	5.7	4209	1	US-08-[illegible]	Sequence [illegible]
c 41	31.8	5.7	4360	1	US-08-[illegible]	Sequence [illegible]
c 42	31.6	5.7	4456	1	US-08-[illegible]	Sequence [illegible]
c 43	31.4	5.7	4356	1	US-08-[illegible]	Sequence [illegible]
c 44	31.4	5.7	4356	1	US-08-[illegible]	Sequence [illegible]
c 45	31.6	5.7	4409	1	US-08-[illegible]	Sequence [illegible]

ALIGNMENTS

RESULT 1
US-08-051-006-5
; Sequence 5, Application US/08051006
; Patent No. 5290924
; GENERAL INFORMATION:
; APPLICANT: Lusk, David L.
; APPLICANT: Brotherton, Richard J.
; APPLICANT: Chamberlain, Douglas A.
; APPLICANT: Larkin, Philip D.
; APPLICANT: Matsu, Ellen S.
; APPLICANT: [illegible] W. James
; APPLICANT: Dennis, Elizabeth S.
; APPLICANT: [illegible] Mark M.
; APPLICANT: Ellis, Jeffrey G.
; TITLE OF INVENTION: A Recombinant Promoter for Gene
; TITLE OF INVENTION: Expression in Monocotyledonous Plants
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Gray, Cary, Ames & Frye
; STREET: 401 B Street, Suite 1
; CITY: San Diego
; STATE: CA
; COUNTRY: USA
; ZIP: 92101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/051,006
; FILING DATE: 19930421
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Brotherton, Harris E.
; REGISTRATION NUMBER: 35,160
; REFERENCE/DOCKET NUMBER: 1000.2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/699-2630
; TELEFAX: 619/236-1048
; TELEX: 695 4271
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 554 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: N





```
Db    589 TGCAAGGAGGCGTTTCTTTCTTTGAATTTAACTAACTCGTTGAGTGGCCCTCTTTCTCGC 648

Qy    428 ACGTAAGGCCTTTCCTCCTCCACACATGTCCATTCGAATTTTACCGTCTTTAGCAAGCGC 487
          |||||||||||||||||||||||||||||||||||||||||||  ||:||||||||||
Db    649 ACGTAAGGCCTTTCCTGCTCCACACATGTCCATTCGAATTTTACCCTCTTTAGCAAGCGC 708

Qy    488 CAAAAGTTTGCATCTTGATGATTTAGCTTGACTATGCGATTGCTTTCCTGGATTCTTGCA 547
          ||||||||||||||||||||||||||||||||||||||||:||||||:|||||  ||.||
Db    709 CAAAAGTTTGCATCTTGATGATTTAGCTTGACTATGCGATTGCTTTCCTGGACCCTTGCA 768

Qy    548 GCT 550
          |||
Db    769 GCT 771
```

RESULT 3
US-09-480-142-2
; Sequence 2, Application US/09480142
; Patent No. 6372963
; GENERAL INFORMATION:
; APPLICANT: D'HALLUIN, Kathleen
; TITLE OF INVENTION: IMPROVED TRANSFORMATION METHOD FOR
; PLANTS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/480,142
; FILING DATE: 10-Jan-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/808,988
; FILING DATE: 20-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Schulman, Robert M.
; REGISTRATION NUMBER: 31,196
; REFERENCE/DOCKET NUMBER: 021565-02
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1086 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "coding region of the bar
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..233
; OTHER INFORMATION: /product= "phosphinothricin
; acetyltransferase (N-terminal half)"
; FEATURE:
; NAME/KEY: intron
; LOCATION: 234..769
; OTHER INFORMATION: /standard_name= "adh1 intron"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 770..1086
; OTHER INFORMATION: /product= "phosphinothricin acetyl
; transferase (c-terminal half)"

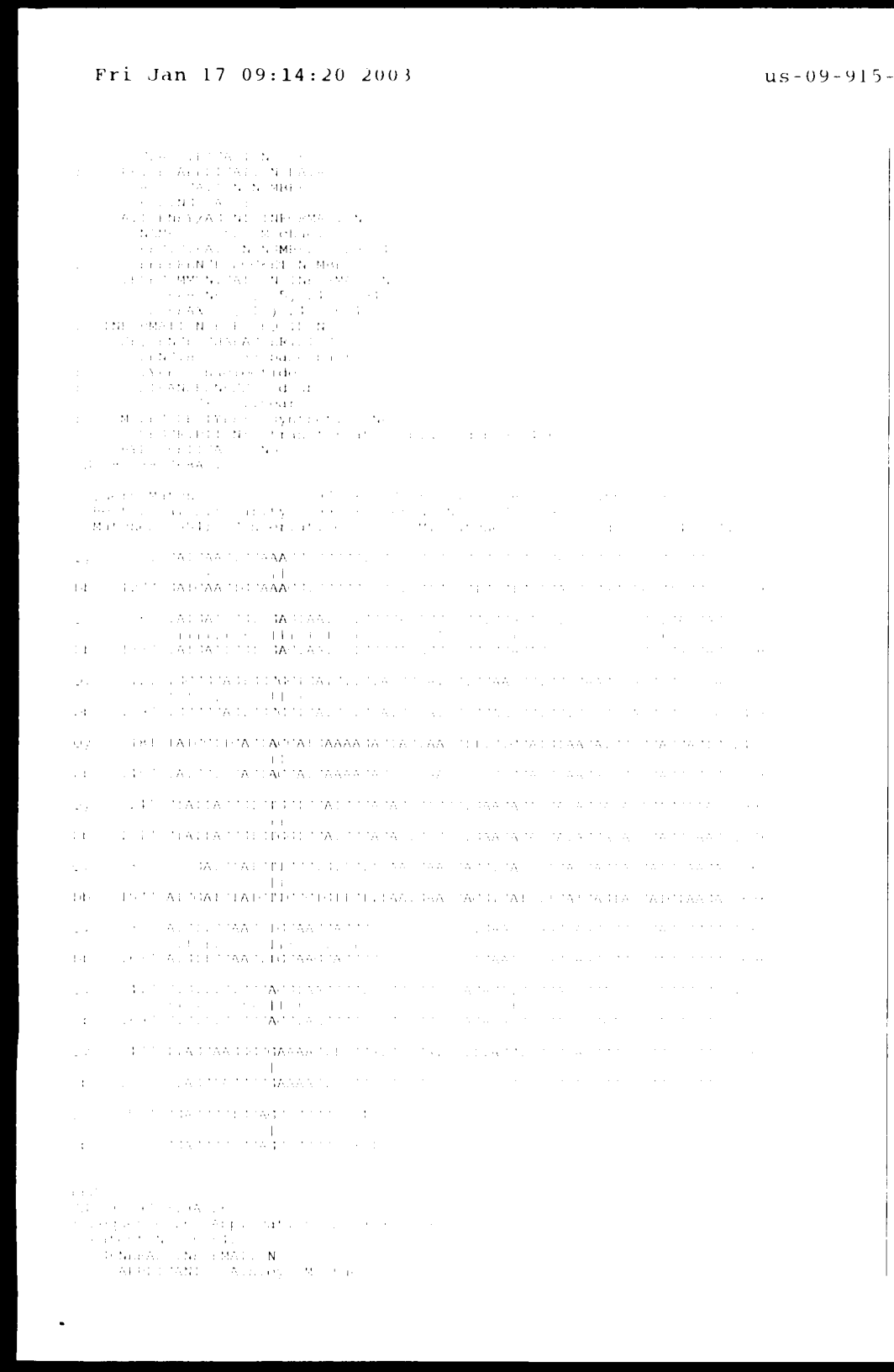
; SEQUENCE DESCRIPTION: SEQ ID NO: [illegible]
US-09-480-142-2

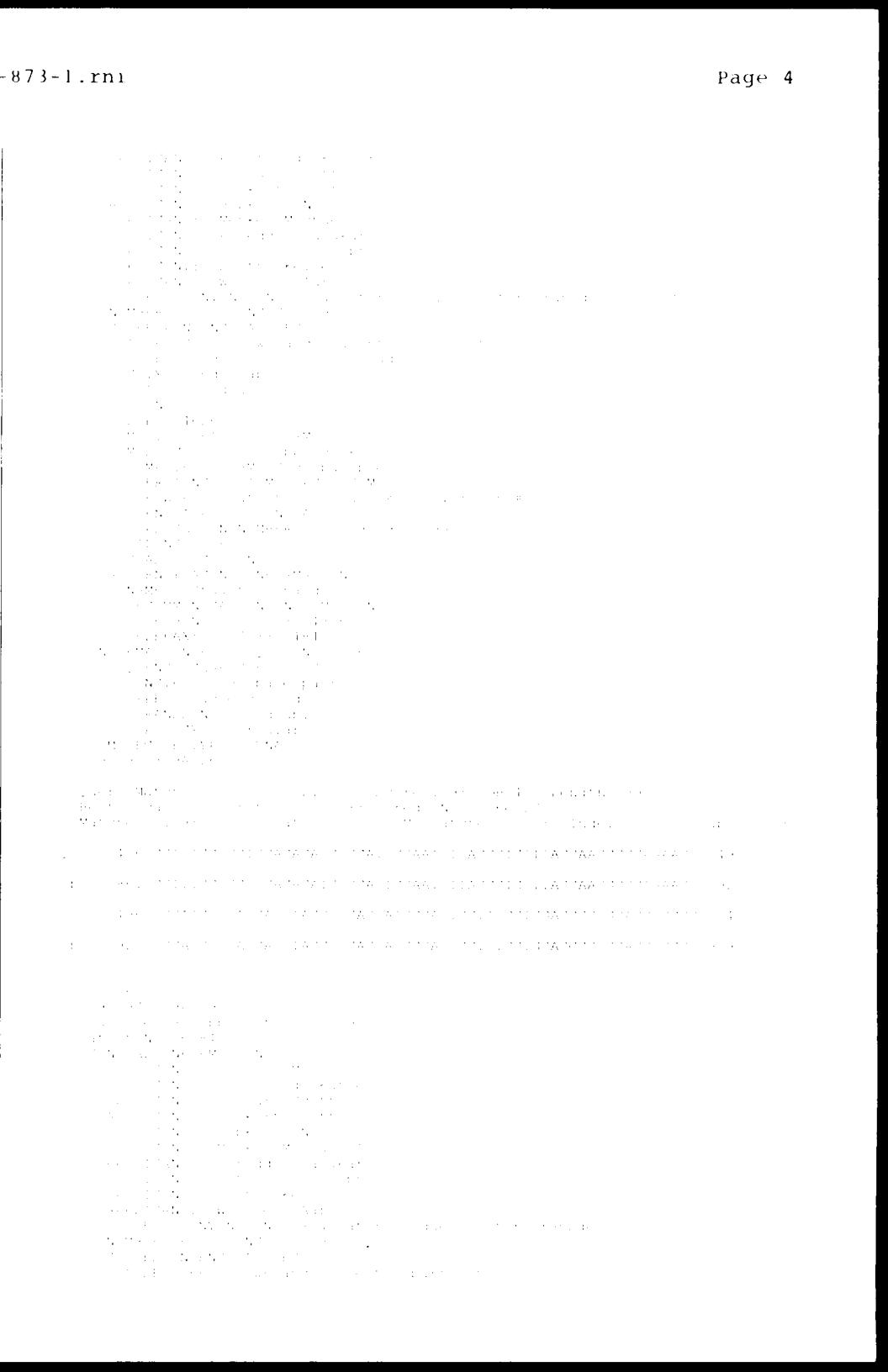
Query Match [illegible]
Best Local Similarity [illegible]
Matches 542; Conservative 0; Mismatches 1; Indels [illegible] Gaps [illegible]

[illegible]

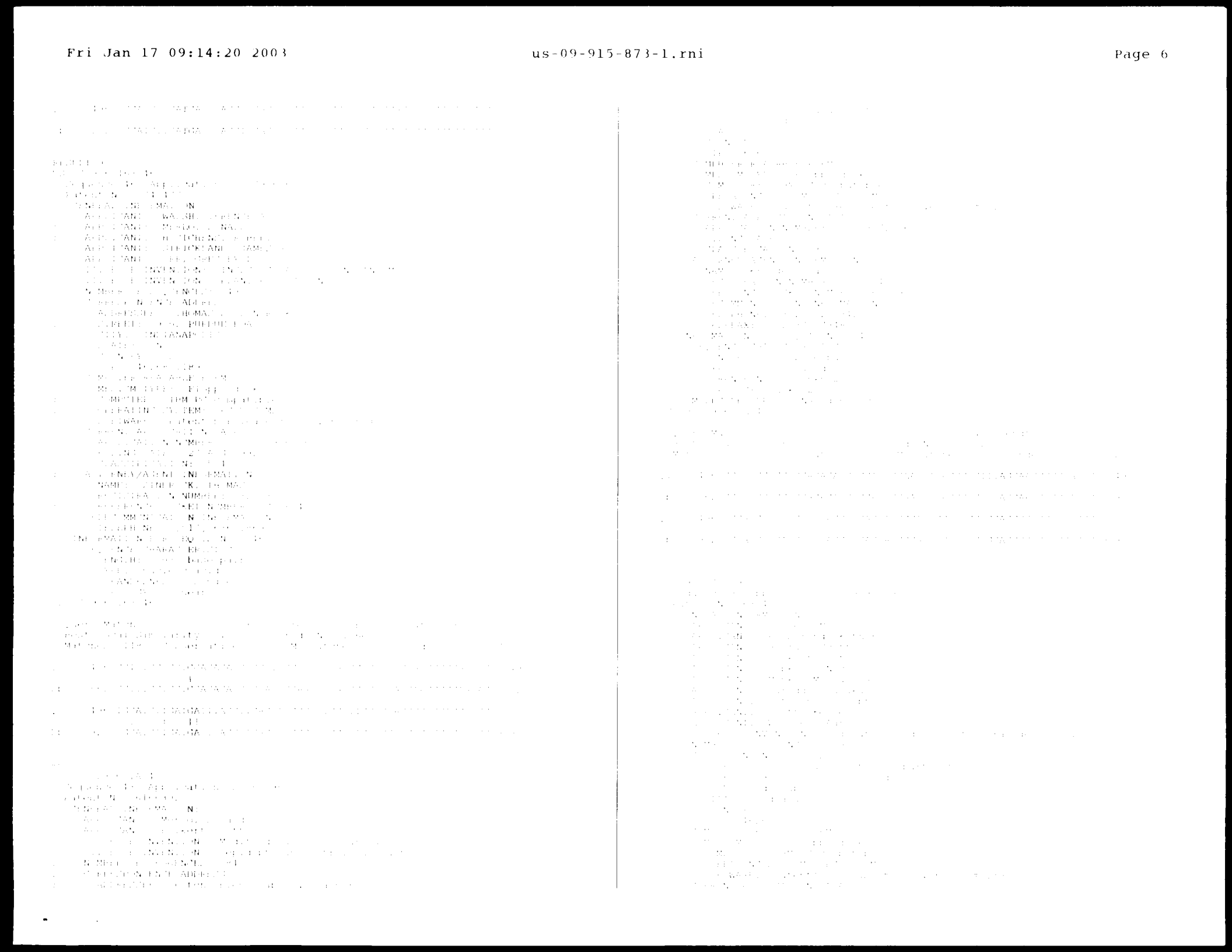
```
Db    769 GCT 771
```

RESULT 4
US-08-838-768A-1
; Sequence 1, Application US/[illegible]
; Patent No. [illegible]
; GENERAL INFORMATION:
; APPLICANT: McMullen, Michael [illegible]
; APPLICANT: [illegible]
; TITLE OF INVENTION: MAIZE [illegible] VIRUS RESISTANCE
; NUMBER OF SEQUENCES: [illegible]
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pioneer Hi-Bred International, Inc.
; STREET: [illegible] Capital Square [illegible]
; STREET: [illegible] Street
; CITY: Des Moines
; STATE: Iowa
; COUNTRY: United States
; ZIP: [illegible]
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, [illegible]
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS, [illegible] Windows
; SOFTWARE: Microsoft Windows [illegible]
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/[illegible]
; FILING DATE: [illegible]





```
;       STREET:  9330 Zionsville Road
;       CITY:  Indianapolis
;       STATE:  Indiana
;       COUNTRY:  USA
;       ZIP:  46268
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/097,419A
;       FILING DATE:
;       CLASSIFICATION:  800
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Stuart, Donald R
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  317 337 4816
;       TELEFAX:  317 337 4847
;   INFORMATION FOR SEQ ID NO:  19:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  9335 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  circular
;     MOLECULE TYPE:  DNA
US-09-097-419A-19

  Query Match             21.5%;  Score 119;  DB 4;  Length 9335;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-27;
  Matches  119;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     436 CCTTTGCTGCTCCACACATGTCCATTCGAATTTACCGTCTTTAGCAACTTCAAAAGTTT 495
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5700 CCTTTGCTGCTCCACACATGTCCATTCGAATTTACCGTCTTTAGCAACTTCAAAAGTTT 5759

Qy     496 TGCATCTTGATGATTTAGCTTGACTATGCCATTGCTTTCCTGGACCCGTCCACTCGAAT 555
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5760 TGCATCTTGATGATTTAGCTTGACTATGCCATTGCTTTCCTGGACCCGTCCACTCGAAT 5819


RESULT 7
US-08-729-601A-47
; Sequence 47, Application US/09729601A
; Patent No. 6166302
;  GENERAL INFORMATION:
;    APPLICANT:  Merlo, Donald J.
;    APPLICANT:  Folkerts, Otto
;    TITLE OF INVENTION:  Modified Bacillus Thuringiensis Gene for
;    TITLE OF INVENTION:  Lepidopteran Control in Plants
;    NUMBER OF SEQUENCES:  84
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fitch, Even, Tabin & Flannery
;      STREET:  135 S. LaSalle St.
;      CITY:  Chicago
;      STATE:  IL
;      COUNTRY:  USA
;      ZIP:  60603
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/729,601A
;      FILING DATE:
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Krueger, James P.
;      REGISTRATION NUMBER:  35,234
;      REFERENCE/DOCKET NUMBER:  60089
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-372-7842
;      TELEFAX:  312-372-7848
;  INFORMATION FOR SEQ ID NO:  47:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  [illegible] base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  DNA (genomic)
US-08-729-601A-47

  Query Match             [illegible]  Score [illegible];  DB 4;  Length [illegible];
  Best Local Similarity   [illegible]  Pred. No. [illegible];
  Matches  [illegible];  Conservative    [illegible];  Mismatches    [illegible];  Indels    [illegible];  Gaps    [illegible];

Qy     [illegible]
          [illegible]
Db     [illegible]

Qy     [illegible]
          [illegible]
Db     [illegible]


RESULT 8
US-08-729-601A-43
; Sequence 43, Application US/09729601A
; Patent No. 6166302
;  GENERAL INFORMATION:
;    APPLICANT:  Merlo, Donald J.
;    APPLICANT:  Folkerts, Otto
;    TITLE OF INVENTION:  Modified Bacillus Thuringiensis Gene for
;    TITLE OF INVENTION:  Lepidopteran Control in Plants
;    NUMBER OF SEQUENCES:  84
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fitch, Even, Tabin & Flannery
;      STREET:  135 S. LaSalle St.
;      CITY:  Chicago
;      STATE:  IL
;      COUNTRY:  USA
;      ZIP:  60603
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/729,601A
;      FILING DATE:
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Krueger, James P.
;      REGISTRATION NUMBER:  35,234
;      REFERENCE/DOCKET NUMBER:  60089
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-372-7842
;      TELEFAX:  312-372-7848
;  INFORMATION FOR SEQ ID NO:  43:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  [illegible] base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  DNA (genomic)
US-08-729-601A-43

  Query Match             [illegible]  Score [illegible];  DB 4;  Length [illegible];
  Best Local Similarity   [illegible]  Pred. No. [illegible];
  Matches  [illegible];  Conservative    [illegible];  Mismatches    [illegible];  Indels    [illegible];  Gaps    [illegible];

Qy     [illegible]
          [illegible]
Db     [illegible]
```



; APPLICATION NUMBER: US/09/097,819A
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Stuart, Donald R
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317 337 4816
; TELEFAX: 317 337 4847
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10160 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA
US-09-097-819A-8

Query Match 21.4%; Score 118; DB 1; Length 1
Best Local Similarity 100.0%; Pred. No. 1.2e-27;
Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 436 GCTTTGCTGCTCCACACATGTCCATTCGAATTTTACTCTTTTACAAAAGGTAAAAAATT 495
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3025 GCTTTGCTGCTCCACACATGTCCATTCGAATTTTACTCTTTTACAAAAGGTAAAAAATT 3084

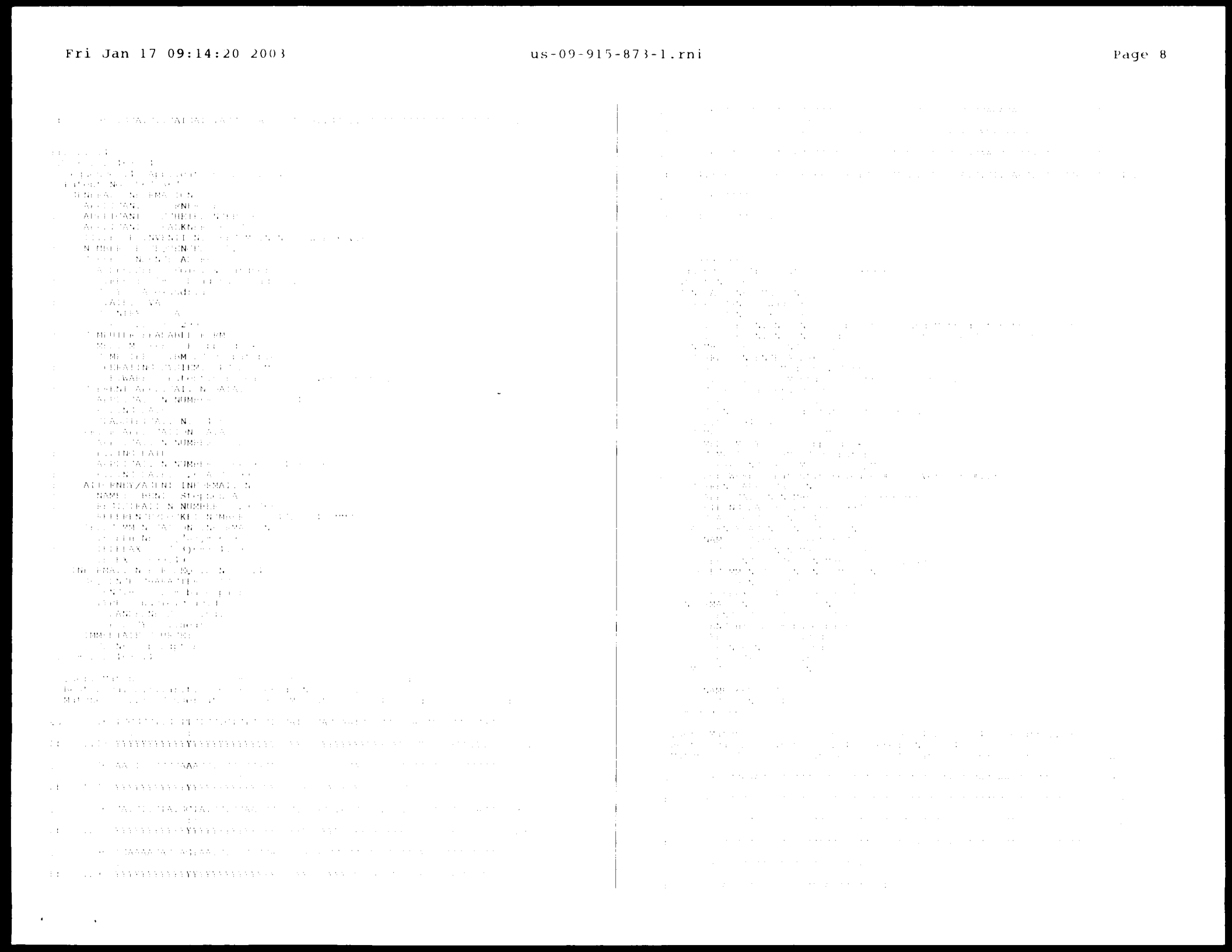
Qy 496 TGCATCTTGATGATTTAGCTTGACTATGCGATTGCTTTCCTGGACCCCTCCAGCTCCTT 555
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3085 TGCATCTTGATGATTTAGCTTGACTATGCGATTGCTTTCCTGGACCCCTCCAGCTCCTT 3144

RESULT 12
US-09-097-819A-9
; Sequence 9, Application US/09/097,819A
; Patent No. 6384207
; GENERAL INFORMATION:
; APPLICANT: Ainley, Michael
; APPLICANT: Armstrong, Katherine
; APPLICANT: Belmar, Scott
; APPLICANT: Folkerts, Otto
; APPLICANT: Hopkins, Nicole
; APPLICANT: Menke, Michael A.
; APPLICANT: Pareddy, Dayaka
; APPLICANT: Petolino, Joseph F.
; APPLICANT: Smith, Kelley
; APPLICANT: Woosley, Aaron
; TITLE OF INVENTION: Regulatory Sequences for Transgenic Plants
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DowElanco Patent Department
; STREET: 9330 Zionsville Road
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: USA
; ZIP: 46268
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/097,819A
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Stuart, Donald R
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317 337 4816
; TELEFAX: 317 337 4847
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11784 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA
US-09-097-819A-9

Query Match
Best Local Similarity
Matches Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy
Db

RESULT 13
US-09-097-819A-10
; Sequence 10, Application US/09/097,819A
; Patent No. 6384207
; GENERAL INFORMATION:
; APPLICANT: Ainley, Michael
; APPLICANT: Armstrong, Katherine
; APPLICANT: Belmar, Scott
; APPLICANT: Folkerts, Otto
; APPLICANT: Hopkins, Nicole
; APPLICANT: Menke, Michael A.
; APPLICANT: Pareddy, Dayaka
; APPLICANT: Petolino, Joseph F.
; APPLICANT: Smith, Kelley
; APPLICANT: Woosley, Aaron
; TITLE OF INVENTION:
; NUMBER OF SEQUENCES:
; CORRESPONDENCE ADDRESS:
; ADDRESSEE:
; STREET:
; CITY:
; STATE:
; COUNTRY:
; ZIP:
; COMPUTER READABLE FORM:
; MEDIUM TYPE:
; COMPUTER:
; OPERATING SYSTEM:
; SOFTWARE:
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE:
; TELEFAX:
; INFORMATION FOR SEQ ID NO:
; SEQUENCE CHARACTERISTICS:
; LENGTH:
; TYPE:
; STRANDEDNESS:
; TOPOLOGY:
; MOLECULE TYPE:



Search completed: January 17, 2003, 05:23:39
Job time : 97 secs